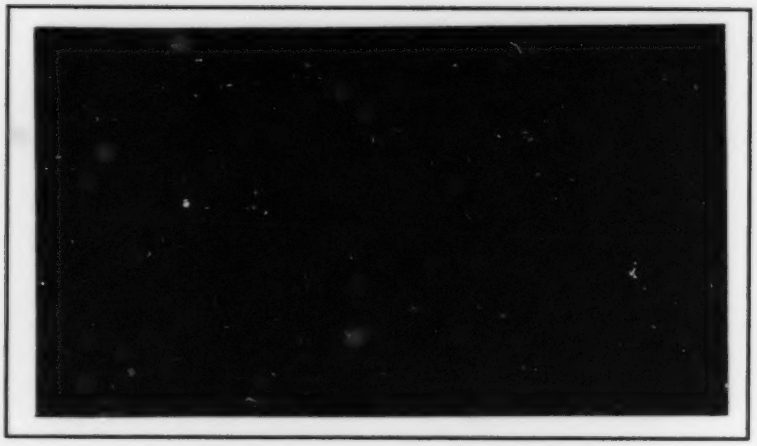

CENTRE FOR
SYSTEMS
SCIENCE

LCCR

Simon Fraser University, Burnaby, B.C., Canada V5A 1S6 / (604) 291-4704

More Songs About Buildings and Food- Collected Extended and Augmented Proofs for a Logic of Default Properties and for Default Reasoning I

James P. Delgrande

CSS/LCCR TR 89-12

November 1989

92-05699/

More Songs About Buildings and Food – Collected, Extended and Augmented Proofs for a Logic of Default Properties and for Default Reasoning I

James P. Delgrande

School of Computing Science,
Simon Fraser University,
Burnaby, B.C.,
Canada V5A 1S6

jim@cs.sfu.ca

Abstract

This technical report will provide a complete and comprehensive collection of proofs concerning a set of logics of default reasoning. It is not intended as an introduction to these systems, and little attempt is made to motivate them; rather the reader is referred to the appropriate papers in the literature. However, the proofs are intended to be complete and comprehensive. Thus I have included all results that have some bearing on later results, or have played some part in the development of the systems. In contrast to conference and journal papers, I have attempted to leave nothing to the imagination: if a partial, non-transparently-obvious result or step in a proof is omitted, then it is a fault of the paper.

1. Introduction

This technical report is intended to provide a complete and comprehensive collection of proofs concerning a set of logics of default properties and for default reasoning. It is not intended as an introduction to these systems, and little attempt is made to motivate them; rather the reader is referred to the appropriate papers in the literature. However, the proofs are intended to be complete and comprehensive. Thus I have included all results that have some bearing on later results, or have played some part in the development of the systems. In contrast to conference and journal papers, I have attempted to leave nothing to the imagination: if a partial, non-transparently-obvious result or step in a proof is omitted, then it is a fault of the paper. I would of course be happy to receive word of any errors or omissions.

Historically, there is the following development of the various systems:

1. A propositional logic for default properties, NP [Delgrande 86a].
2. A first-order logic for default properties, N [Delgrande 86b; Delgrande 87b].
3. Default reasoning with the logic N [Delgrande 87a,c; Delgrande 88].
4. Extending N to a stronger (in fact "the strongest") logic for default properties N+; extending default reasoning to apply in N+.

This last point, the extension to N+, has not appeared in the literature, but presumably and hopefully will do so in the not-too-distant future. If the reader is able to follow the other material, then hopefully s/he will be able to follow the abbreviated development of N+.

Proofs in this report then cover points 1.-4. above. Some of the proofs (notably of soundness and completeness) have appeared in previous reports; others (notably the equivalence of the two approaches to default reasoning) have been substantially reworked and clarified. In addition, the relation between the two semantic approaches (one based on an accessibility relation between worlds, the other based on a worlds-selecting function) has hopefully been clarified somewhat (particularly with the introduction of the systems NP+ and N+). The addition of "I" to the report title would seem to indicate that there is another of these collections in preparation, or at least intended. That indeed is the case and, in this sequel, I hope to include proofs for further work and extensions in this area. This second report is intended to cover the following areas:

5a. A (sometimes) tractable algorithm for default reasoning, based on the (extended) propositional logic NP+.

5b The algorithm of 5a, extended to the full logic N+.

6. Default reasoning in the full class of normal conditional logics.

Of these areas, 5a has been completed, and will be submitted for publication shortly; 5b involves substantially rearranging the algorithm, but appears to present no great new difficulty. (Essentially, for 5a, in propositional default reasoning, the algorithm can be split into three distinct and independent steps; for first-

order reasoning (to avoid certain infinite structures) this is not always the case.) The last area strikes me as being extremely interesting, and involves extending the notion of default reasoning to deal with counterfactual statements, statements of obligation, and so on. I hope to have something prepared in that area in time for the 1990 AAAI.

The logics presented here are very closely related. For the original formulation, the semantics was motivated in terms of a possible worlds semantics, incorporating a standard accessibility relation between possible worlds. However the semantics was actually specified using a "world-selection" function instead of an accessibility relation; consequently it was necessary to show that the world-selection function conformed to the properties of the accessibility relation (see [Delgrande 86b]). The resulting systems, NP and N, based on the world-selection function appeared (well, to my mind) to adequately capture the requisite range of default reasoning and, in particular, was at least as expressive as comparable systems. However, one thing that I did not do was to show that the semantic approaches were in any sense equivalent. Now, it turns out that these approaches *aren't* equivalent, but by introducing certain reasonable assumptions can be made so. The resultant strengthening of the world-selection function then yields a function that cannot be further strengthened without violating a condition on the accessibility relation. The resulting logics based on this augmented function, NP+ and N+, are then stronger than the previous logics; however this strengthening manifests itself mainly in reasoning about contrapositives – a form of reasoning that has generally been little studied in default reasoning. Interestingly, the original systems (NP and N) are perfectly adequate for general reasoning about contrapositive information (and in fact I would argue that the original systems are stronger than extant systems with respect to contrapositive reasoning). The strengthening (to NP+ and N+) allows reasoning about contrapositives where negations of default assertions are employed, and then only in quite obscure circumstances (see section 2.5 (Some More Results)). However, the strengthened system is required for the algorithm mentioned in area 5a above.

In sum we have the following systems:

1. NP – The "original" propositional system for expressing default properties.
2. NP+ – The "strongest possible" propositional system.
3. N – The "original" first-order system for expressing default properties.
4. N+ – The "strongest possible" first-order system.

This report is composed of three major sections; they are:

1. The propositional systems NP and NP+.
2. The first-order systems N and NP+.
3. Default reasoning in the various systems.

2. The Propositional Logics

2.1. Preliminary Concerns

The logics presented here are intended to be applicable to assertions of default properties. In the logics "birds fly", for example, is interpreted as "all other things being equal, birds fly", or "*ceteris paribus*, birds fly", or "ignoring exceptional conditions, birds fly". The general idea is that an operator $\Rightarrow$ is introduced into, first, standard propositional logic (PC), and then (in section 3) into the standard first-order predicate calculus (FOL). In either case, the statement $A \Rightarrow B$ is interpreted as "in the normal course of events, if A then B". The logics allow the consistent assertion, for example, that

$$(\forall x)(Raven(x) \Rightarrow Black(x))$$

along with

$$(\forall x)((Raven(x) \wedge Albino(x)) \Rightarrow \neg Black(x)),$$

or, informally, "ravens are normally black" but "albino ravens are normally not black". Intuitively, the connection between the antecedent and consequent of these conditionals may be thought of as a relationship in some scientific theory. Ravens are black, but only if some set of presumed "additional assumptions" is satisfied: the raven isn't albino, hasn't been painted, isn't in a strong red light, etc. More generally, the approach is intended also to be applicable to general statements of typicality, and so we can use this connective to represent statements such as "typically Quakers are pacifists" or "typically adults are employed".

In extending the semantics of PC to account for $\Rightarrow$, I initially adopt a possible worlds approach, where $A \Rightarrow B$ is true if B is true in the "least exceptional" worlds where A is true. Roughly then this says that ravens are black, if we "factor out" all exceptional circumstances such as albinoism, being painted, being featherless, etc. While the actual choice of a metric for "less exceptional than" between possible worlds is, to a large extent, a pragmatic affair, we can specify some bounds on such a metric, and these bounds will constrain the semantics of the $\Rightarrow$ operator. The resulting formal system is a "conditional logic", of a class of logics that have been developed for dealing with counterfactual conditionals, conditional obligation, and other such areas. Thus while the system described here differs in significant respects from other conditional logics described in the literature, nonetheless the semantic theory presented is in much the same spirit as these other logics.

Hence at the outset I consider a possible worlds formulation, where the accessibility relation E is interpreted so that Ew_1w_2 holds between worlds w_1 and w_2 just when w_2 is at least as uniform, or at least as unexceptional, as w_1. From this $A \Rightarrow B$ is defined to be true at a world just when the least exceptional worlds in which A is true also have B true, and in no less exceptional world is A true and B false.

The following conditions are argued (in [Delgrande 87b]) to be sufficient to capture our informal reading of E as obtaining between relatively "less exceptional" worlds.

Reflexive: Eww for all worlds w.

Transitive: If Ew_1w_2 and Ew_2w_3 then Ew_1w_3.

Forward Connected: If Ew_1w_2 and Ew_1w_3 then either Ew_2w_3 or Ew_3w_2.

That is, every world is at least as unexceptional as itself, and the notion of unexceptionalness is transitive. Note that the last condition is weaker than the euclidean condition found, for example, in S5 where the "either $\cdots$ or $\cdots$" is replaced by "$\cdots$ and $\cdots$".

With these conditions, the accessibility relation for a given world w yields a well-ordering on mutually disjoint sets of worlds, with w in the "topmost" set. If only a finite number of worlds are considered, then there is a set of simplest worlds, all of which are mutually accessible. Viewed "globally" the accessibility relation imposes a backwards tree structure on sets of worlds. Two worlds may then be mutually inaccessible, yet have common accessible worlds. That is, we may have Ew_1w_3 and Ew_2w_3 but neither Ew_1w_2 nor Ew_2w_1. The modal logic corresponding to this accessibility relation is the standard temporal logic S4.3; it subsumes S4 but does not subsume S5.

However, having motivated a formal semantics for an accessibility relation E, it proved to be the case that the technical development is considerably simplified if we consider not a semantics based on this relation E, but rather deal with a *world-selection* function f. Accessibility is replaced by a function which, given a world and a proposition, returns the set of the least exceptional worlds in which the proposition is true. If we identify propositions with sets of possible worlds and write $\|A\|^M$ to be the set of worlds in which A is true in model M then, for some possible state of affairs w, $f(w, \|A\|^M)$ is the set of least exceptional states of affairs in which A is true. Thus $A \Rightarrow B$ is true at w if and only if, for each of the least exceptional states of affairs in which A is true, B is also true. This means that $A \Rightarrow B$ is true at w if and only if $f(w, \|A\|^M) \subseteq \|B\|^M$. The problem now of course is to specify conditions on f that would obtain in a model where the accessibility relation is reflexive, transitive, and forward connected. This issue is taken up next.

2.2. A Formal Semantics

The propositional language L_p I shall consider is formed in the usual manner from a denumerable set of atomic sentences $\mathbf{P} = \{p_0, p_1, \cdots\}$, together with standard connectives $\neg$ and $\supset$ for negation and the material conditional, and a new binary connective $\Rightarrow$ for the variable conditional. However we do not allow an occurrence of $\Rightarrow$ to appear within the scope of (i.e. in an argument of) another $\Rightarrow$ operator.

Definition 1: The *sentences* or *well-formed formulae (wffs)* of L_p comprise the least set such that:

(i) If $p_i \in \mathbf{P}$ then p_i is a wff.

(ii) If A, B are wffs then $\neg A$, $(A \supset B)$, $(A \Rightarrow B)$, are wffs.

Conjunction ($\wedge$), disjunction ($\vee$), and biconditionality ($\equiv$) are introduced by definition. The letters $A, B, C, \cdots$ will stand for arbitrary well-formed formulae of L_p. If A is a wff of L_p, we may write $A \in L_p$.

Sentences of L_p are interpreted in terms of a *model* $M = \langle W, f, P \rangle$ where W is a set, f is a function from $W \times 2^W$ to 2^W and P is a function from P to 2^W. Informally, W is a set of possible worlds. For wff A, the symbolism $\|A\|^M$ will be used to stand for the set of worlds in M in which A is true. If we identify propositions with possible worlds, then $\|A\|^M$ stands for the proposition expressed by A. The function f then picks out a set of possible worlds $f(w, \|A\|^M)$ for each possible world w and proposition $\|A\|^M$. P maps atomic sentences onto those worlds where the sentence is true.

The symbolism $\models^M_w A$ is used to assert that A is *true in the model* M *at world* w. If some M and w are understood, we will simply say that A is *true*. We write $\models A$ in the case that A is true at every world in every model, and say that A is *valid*.[1] A is *satisfiable* if and only if $\neg A$ is not valid. Truth at a world w in a model $M = \langle W, f, P \rangle$ is defined by:

Definition 2:

(i) $\models^M_w p_n$ iff $w \in P(p_n)$ for every $n \geq 0$.

(ii) $\models^M_w \neg A$ iff not $\models^M_w A$.

(iii) $\models^M_w A \supset B$ iff if $\models^M_w A$ then $\models^M_w B$.

(iv) $\models^M_w A \Rightarrow B$ iff $f(w, \|A\|^M) \subseteq \|B\|^M$.

Definition 3: By a *NP-model*, I will mean a model $M = \langle W, f, P \rangle$ where the following conditions on f are met:

(i) $f(w, \|A\|^M) \subseteq \|A\|^M$.

(ii) If $f(w, \|A\|^M) \subseteq \|B\|^M$ then $f(w, \|A\|^M) \subseteq f(w, \|A \wedge B\|^M)$.

(iii) If $f(w, \|A\|^M) \not\subseteq \|B\|^M$ then $f(w, \|A \wedge \neg B\|^M) \subseteq f(w, \|A\|^M)$.

(iv) $f(w, \|A \vee B\|^M) \subseteq f(w, \|A\|^M) \cup f(w, \|B\|^M)$.

As an aside, note that $f(w, \|A\|^M) = \varnothing$ does not entail that $\|A\|^M = \varnothing$. However $f(w, \|A\|^M) = \varnothing$ does entail that $f(w, \|B\|^M) \cap \|A\|^M = \varnothing$ for every $B \in L_p$.

Definition 4: By a *NP+-model*, I will mean a model $M = \langle W, f, P \rangle$ where the additional conditions on f are met:

(v) If $f(w, \|\neg A\|^M) = \varnothing$ then $\models^M_w A$.

---

[1] This notation is a little unfortunate, since validity (and other aspects of truth) will be with respect to either NP-models or NP+-models. The context of the exposition however will make clear which set of models we are talking about.

(vi) If $f(w, \|A\|^M) \subseteq \|C\|^M$ and $f(w, \|A \vee B\|^M) \not\subseteq \|C\|^M$ then $f(w, \|\neg C\|^M) \subseteq \|\neg A\|^M$.

Note that if we can order sets of worlds (using $\leq$ for notation – see section 5) this last condition can be equivalently expressed as:

(vi') If $f(w, \|A\|^M) \subseteq \|C\|^M$ and $f(w, \|\neg C\|^M) \leq f(w, \|A\|^M)$ then
$f(w, \|\neg C\|^M) \subseteq \|\neg A\|^M$.

Theorem 1: If $f(w, \|A\|^M) \subseteq \|B\|^M$ and $f(w, \|B\|^M) \subseteq \|A\|^M$ then $f(w, \|A\|^M) = f(w, \|B\|^M)$.

Proof:

If $f(w, \|A\|^M) \subseteq \|B\|^M$ then $f(w, \|A\|^M) \not\subseteq \|\neg B\|^M$. (Assuming clearly that $f(w, \|A\|^M) \neq \emptyset$ – otherwise there is no problem.)

Hence by condition (iii), $f(w, \|A \wedge B\|^M) \subseteq f(w, \|A\|^M)$.

Also, since $f(w, \|A\|^M) \subseteq \|B\|^M$ then by condition (ii), $f(w, \|A\|^M) \subseteq f(w, \|A \wedge B\|^M)$.

Thus if $f(w, \|A\|^M) \subseteq \|B\|^M$ then $f(w, \|A\|^M) = f(w, \|A \wedge B\|^M)$.

Similarly if $f(w, \|B\|^M) \subseteq \|A\|^M$ then $f(w, \|B\|^M) = f(w, \|A \wedge B\|^M)$.

So if $f(w, \|A\|^M) \subseteq \|B\|^M$ and $f(w, \|B\|^M) \subseteq \|A\|^M$ then by transitivity of $=$, $f(w, \|A\|^M) = f(w, \|B\|^M)$.

Theorem 2:

$f(w, \|A \vee B\|^M) = f(w, \|A\|^M)$ or

$f(w, \|A \vee B\|^M) = f(w, \|B\|^M)$ or

$f(w, \|A \vee B\|^M) = f(w, \|A\|^M) \cup f(w, \|B\|^M)$.

Proof:

(i) Assume that $f(w, \|A \vee B\|^M) \cap \|A\|^M = \emptyset$.

Since $f(w, \|A \vee B\|^M) \subseteq \|A \vee B\|^M$ we obtain that $f(w, \|A \vee B\|^M) \subseteq \|B\|^M$.

Since $f(w, \|B\|^M) \subseteq \|B\|^M$ and so $f(w, \|B\|^M) \subseteq \|A \vee B\|^M$ we get via the above theorem that $f(w, \|A \vee B\|^M) = f(w, \|B\|^M)$ which is the second disjunct of the result to be shown.

(ii) Similarly if we assume that $f(w, \|A \vee B\|^M) \cap \|B\|^M = \emptyset$ we obtain that $f(w, \|A \vee B\|^M) = f(w, \|A\|^M)$ which is the first disjunct of the result to be shown.

(iii) So we can assume that $f(w, \|A \vee B\|^M) \cap \|A\|^M \neq \emptyset$ and $f(w, \|A \vee B\|^M) \cap \|B\|^M \neq \emptyset$.

For the first condition we obtain that $f(w, \|A \vee B\|^M) \not\subseteq \|\neg A\|^M$.

Condition (iii) on f then yields $f(w, \|(A \vee B) \wedge A\|^M) \subseteq f(w, \|A \vee B\|^M)$ and so $f(w, \|A\|^M) \subseteq f(w, \|A \vee B\|^M)$.

Similarly for the second condition we obtain $f(w, \|B\|^M) \subseteq f(w, \|A \vee B\|^M)$, and so $f(w, \|A\|^M) \cup f(w, \|B\|^M) \subseteq f(w, \|A \vee B\|^M)$.

This together with condition (iv) give the third disjunct of the result to be shown, and hence the required result.

Note that this result is obtained using conditions i-iv only.

We can tie this semantics to our previous informal development, in terms of an accessibility relation between possible worlds, as follows. An E-model is defined as a structure $M = <W, E, P>$ where W and P are as before and E is a reflexive, transitive, forward connected relation on W. Truth in an E-model is defined as expected; the only interesting case is for $A \Rightarrow B$ wherein

$A \Rightarrow B$ is true in M at a world w iff

(i) there is a $w_1 \in W$ so that Eww_1 and $\vdash^M_{w_1} A$ and $\vdash^M_{w_1} B$

and for every $w_2 \in W$ where Ew_1w_2, $\vdash^M_{w_1} A \supset B$, or

(ii) for every w_1 such that Eww_1, $\vdash^M_{w_1} \neg A$.

That is, there is a world w_1 in which A and B are true, and in no less exceptional worlds is A true and B false. Case (ii) takes care of the situation where A is false in all worlds accessible from w.

The two semantic approaches are not equivalent. Consider first an E-model where A and B are true at all worlds, but where there is no least element in the partial ordering of worlds (so there is an infinite number of worlds). Clearly $\vdash^M_w A \Rightarrow B$ is true at all worlds; equally clearly, since there is no least world where A is true, the corresponding $f(w, \|A\|^M)$ is undefined. (The logics do have the finite model property, but this doesn't address the point at hand.) Hence I adopt the *limit assumption* of Stalnaker [Stalnaker 68]: that if there are worlds where A is true, then there is a set of least worlds where A is true.

Conversely, note that if the function f were defined in the obvious way from the accessibility relation E, then we would have that if $w_1 \in f(w, \|A\|^M)$ then $f(w_1, \|B\|^M) \subseteq f(w, \|B\|^M)$ for every $B \in L_P$. This clearly is not specified in the conditions for f, where the values of f for two worlds (i.e. for $f(w_1, \|B\|^M)$ and $f(w, \|B\|^M)$) are completely independent. However, this only has an effect on the truth of formulae if we allow nested occurrences of the $\Rightarrow$ operator – which, here, we do not. (However, see the later discussion for more details.) Thus, given these restrictions it is reasonable to suppose that the semantic approaches coincide. This in fact is the case; we obtain:

Definition 5: $f_E(w, \|A\|^M) = \{w_1 \mid Eww_1$ and $\vdash^{M'}_{w_1} A$, and for all w_2 such that Ew_1w_2 and $\vdash^{M'}_{w_1} A$, we also have $Ew_2w_1\}$

Theorem 3:

1. If $\vdash^M_w A \supset B$ and $w_1 \in f_E(w, \|A\|^M)$ and $w_2 \in f_E(w, \|B\|^M)$ then Ew_1w_2.
2. If $\vdash^M_w A \Rightarrow B$ and $w_1 \in f_E(w, \|A\|^M)$ and $w_2 \in f_E(w, \|B\|^M)$ then Ew_1w_2.
3. If $\vdash^M_w A \supset B$ and $f_E(w, \|A\|^M) \not\subset f_E(w, \|B\|^M)$ then $f_E(w, \|A\|^M) \cap f_E(w, \|B\|^M) = \emptyset$.

Proof:

1. If $w_1 \in f_E(w, \|A\|^M)$ and $w_2 \in f_E(w, \|B\|^M)$ then Eww_1 and Eww_2.
   Since E is forward connected, Ew_1w_2 or Ew_2w_1.
   The former case is what is to be proved, so assume the latter.
   But $\vdash^M_{w_1} B$ (since $A \supset B$), and so by the definition of f_E, we must have $w_1 \in f_E(w, \|B\|^M)$, and so also Ew_1w_2.

2. Proof is analogous to the first part: If $w_1 \in f_E(w, \|A\|^M)$ and $w_2 \in f_E(w, \|B\|^M)$ then Eww_1 and Eww_2.
   Since E is forward connected, Ew_1w_2 or Ew_2w_1.
   The former case is what is to be proved, so assume the latter.
   But $\vdash^M_{w_1} B$ (since $A \Rightarrow B$), and so by the definition of f_E, we must have $w_1 \in f_E(w, \|B\|^M)$, and so also Ew_1w_2.

3. Assume that $w_1 \in f_E(w, \|A\|^M)$ and $w_1 \notin f_E(w, \|B\|^M)$.
   Assume further that $w_2 \in f_E(w, \|A\|^M)$ and $w_2 \in f_E(w, \|B\|^M)$.
   From part 1. above we have that Ew_1w_2.
   Since $w_1 \in f_E(w, \|A\|^M)$ we also have (from the definition of f_E) that Ew_2w_1.
   However since $\vdash^M_{w_1} B$ we also have (again, from the definition) $w_1 \in f_E(w, \|B\|^M)$.
   But this contradicts the assumption above.
   Hence there is no w_2 so that $w_2 \in f_E(w, \|A\|^M)$ and $w_2 \in f_E(w, \|B\|^M)$.

Theorem 4: If $M' = \langle W, E, P \rangle$ is an E-model satisfying the limit assumption then $M = \langle W, f_E, P \rangle$ is a $NP+$-model.

Proof: We need to show that f satisfies the conditions for an $NP+$-model.

(i): If $w_1 \in f(w, \|A\|^M)$ then $\vdash^M_{w_1} A$,
hence $w_1 \in \|A\|^M$.

(ii): Assume $f(w, \|A\|^M) \neq \emptyset$. (Otherwise we are done.)
Assume also $f(w, \|A\|^M) \subseteq \|B\|^M$.
If $w_1 \in f(w, \|A\|^M)$ then $w_1 \in \|A\|^M$.
Since $w_1 \in \|B\|^M$ also, so $w_1 \in \|A \wedge B\|^M$.
For w_2 where Ew_1w_2 and $\vdash^M_{w_2} A$, then Ew_2w_1, by the definition of f.
But $w_2 \in f(w, \|A\|^M)$ (by transitivity), and so $w_2 \in \|B\|^M$, $\vdash^M_{w_2} A \wedge B$, and $w_1 \in f(w, \|A \wedge B\|^M)$.

(iii): For this condition I prove the contrapositive.
So assume that $f(w, \|A \wedge \neg B\|^M) \not\subseteq f(w, \|A\|^M)$.
So, there is $w_1 \in f(w, \|A \wedge \neg B\|^M)$ such that $w_1 \notin f(w, \|A\|^M)$;
so $\vdash^M_{w_1} A$ and $\vdash^M_{w_1} \neg B$.

In addition, since $(A \wedge \neg B) \supset A$, by the third part of the preceding theorem we have that $f(w, \|A \wedge \neg B\|^M) \cap f(w, \|A\|^M) = \emptyset$. (This last fact is what we will contradict for the proof.)

Assume $w_2 \in f(w, \|A\|^M)$; hence $\vdash^M_{w_2} A$.

Assume without loss of generality that $\vdash^M_{w_2} \neg B$. (If there is no such w_2 then we have that $f(w, \|A\|^M) \subseteq \|B\|^M$ and so would be done.)

We have that Ew_1w_2 from the preceding theorem. Since $\vdash^M_{w_2} A \wedge \neg B$, we must also have Ew_2w_1 (from the definition of F_E).

Now, we must have that $w_2 \in f(w, \|A \wedge \neg B\|^M)$:

> For any w_3 where Ew_2w_3 and $\vdash^M_{w_3} A \wedge \neg B$, we have Ew_1w_3 (from transitivity of E).
>
> Hence Ew_3w_1 (from the definition of f_E).
>
> Hence Ew_3w_2 (by transitivity of E).
>
> Thus by the definition of f_E, $w_2 \in f(w, \|A \wedge \neg B\|^M)$.

But this means that $w_2 \in f(w, \|A \wedge \neg B\|^M)$ and $w_2 \in f(w, \|A\|^M)$, contradicting the original assumption.

Hence if $w_2 \in f(w, \|A\|^M)$ we must have $\vdash^M_{w_2} B$ and so $f_E(w, \|A\|^M) \subseteq \|B\|^M$.

(iv): Let $w_1 \in f(w, \|A \vee B\|^M)$.

So, $\vdash^M_{w_1} A \vee B$ and for any w_2 such that Ew_1w_2 and $\vdash^M_{w_2} A \vee B$ then Ew_2w_1.

Assume that $\vdash^M_{w_1} A$.

So for w_2 such that Ew_1w_2 and $\vdash^M_{w_2} A \vee B$ and $\vdash^M_{w_2} A$ then Ew_2w_1.

So, $w_1 \in f(w, \|A\|^M)$.

By a similar argument for B the assertion is proved.

(v): If $f_E(w, \|\neg A\|^M) = \emptyset$ then by definition of f_E, $\{w_1 \mid Eww_1$ and $\vdash^{M'}_{w_1} \neg A$, and for all w_2 such that Ew_1w_2 and $\vdash^{M'}_{w_2} \neg A$, we also have $Ew_2w_1\} = \emptyset$.

We have that Eww, and so from the definition this means that if $\vdash^{M'}_{w} \neg A$ that, one way or another (i.e., the proof is trivial) $f_E(w, \|\neg A\|^M) \neq \emptyset$.

Hence it must be that not $\vdash^M_w \neg A$ and so $\vdash^M_w A$.

(vi): Assume that $w_1 \in f_E(w, \|A\|^M)$ and $w_2 \in f_E(w, \|A \vee B\|^M)$.

Hence by the previous theorem, Ew_1w_2.

Since $f_E(w, \|A \vee B\|^M) \not\subseteq \|C\|^M$ assume without loss of generality that $\vdash^M_{w_2} \neg C$; hence $\vdash^M_{w_2} \neg A$.

(Otherwise if $\vdash^M_{w_2} A$ then Ew_2w_1 and (eventually) $w_2 \in f_E(w, \|A\|^M)$, which is a contradiction.)

If for every $w_2 \in f_E(w, \|\neg C\|^M)$ we have $w_2 \in \|\neg A\|^M$ then we're done.

So assume $w_2 \notin f(w, \|\neg C\|^M)$ and let $w_3 \notin f(w, \|\neg C\|^M)$.

We must have one of Ew_2w_3 or Ew_3w_2. Since Ew_3w_2 gives that $w_2 \in f(w, \|\neg C\|^M)$ we must

have Ew_2w_3.

If $\vdash_{w_3}^{M} A$ then, since Ew_1w_3 (via transitivity of E), we get that $w_3 \in f(w, \|A\|^{M})$, contradicting $f(w, \|A\|^{M}) \subseteq \|C\|^{M}$.

Hence $\vdash_{w_3}^{M} \neg A$ and so $f(w, \|\neg C\|^{M}) \subseteq \|\neg A\|^{M}$.

Hence for every E-model satisfying the limit assumption, there is a corresponding $NP+$-model. However, we need one further step, and that is to show that these models verify exactly the same sentences at each possible world. This is accomplished in the following theorem.

Theorem 5: If $M_E = <W, E, P>$ is an E-model satisfying the limit assumption then there is a $NP+$-model $M_f = <W, f_E, P>$ such that M_E and M_f verify exactly the same sentences. That is, for $A \in L_P$, $\vdash_{w}^{M_E} A$ iff $\vdash_{w}^{M'} A$.

Proof: Given M_E, construct M_f by defining f_E as in definition 5. The proof is trivial for $\Rightarrow$-free formulae.

For $A \Rightarrow B \in L_P$ we obtain:

$\vdash_{w}^{M_E} A \Rightarrow B$ iff

(a) there is a $w_1 \in W$ such that Eww_1 and $\vdash_{w_1}^{M_E} A$ implies $\vdash_{w_1}^{M_E} B$ and for every $w_2 \in W$ where Ew_1w_2, $\vdash_{w_2}^{M_E} A \supset B$ or

(b) for every w_1 such that Eww_1, $\vdash_{w_1}^{M_E} \neg A$.

Case (b) implies $f(w, \|A\|^{M'}) = \varnothing$ and so $f(w_1, \|A\|^{M'}) \subseteq \|B\|^{M'}$ for all B.

Case (a), together with the limit assumption, guarantees that $f(w, \|A\|^{M'}) \neq \varnothing$.

For $w_1 \in f(w, \|A\|^{M'})$, $\vdash_{w_1}^{M'} A$.

Also, since $\vdash_{w_1}^{M_E} A \supset B$, an inductive step (on modality-free formulae) gives $\vdash_{w_1}^{M'} A \supset B$.

Hence by modus ponens, $\vdash_{w_1}^{M'} B$ and so $w_1 \in \|B\|^{M'}$.

Thus $f(w, \|A\|^{M'}) \subseteq \|B\|^{M'}$.

The proof for not $\vdash_{w}^{M_E} A \Rightarrow B$ proceeds analogously, and is omitted.

Hence $\vdash_{w}^{M_E} A$ iff $\vdash_{w}^{M'} A$.

The world selection function f though, as specified, is too weak to allow the binary relation E to be recovered from it. Basically, f is adequate only for reasoning with formulae where no conditional operator appears within the scope of another. That is, none of the conditions on f deal with nested or iterated occurrences of the conditional. Since the language is restricted to disallow nested occurrences of the $\Rightarrow$ operator, this presents no problem.

The following lemma has an obvious correlate with the second theorem of this section. I proved the two about a year or so apart and so didn't exploit the redundancy (in proving the previous theorem); at present I can't be bothered having one be a theorem and the other a straightforward corollary. Anyway, if anyone is actually reading this, it's safe to skip the lemma. We obtain:

Lemma 1: $f(w,\|A\|^M)=f(w,\|A\wedge B\|^M)$ or $f(w,\|A\|^M)=f(w,\|A\wedge\neg B\|^M)$ or

$$f(w,\|A\|^M)=f(w,\|A\wedge B\|^M)\cup f(w,\|A\wedge\neg B\|^M).$$

Proof: Assume $f(w,\|A\|^M)\neq\varnothing$. (Otherwise we're done.)

By the law of the excluded middle for PC, we have:

$$(A\Rightarrow B\vee A\Rightarrow\neg B)\vee(\neg(A\Rightarrow B\vee A\Rightarrow\neg B)).$$

By the truth conditions for $\Rightarrow$ we have that $A\Rightarrow B$ is true iff $f(w,\|A\|^M)\subseteq\|B\|^M$.

Hence, from the instance of the law of the excluded middle above,

$$f(w,\|A\|^M)\subseteq\|B\|^M \text{ or } f(w,\|A\|^M)\subseteq\|\neg B\|^M) \text{ or}$$

$$f(w,\|A\|^M)\not\subseteq\|B\|^M \text{ and } f(w,\|A\|^M)\not\subseteq\|\neg B\|^M. \tag{1}$$

Condition (ii) in the definition of f states that:

$$\text{if } f(w,\|A\|^M)\subseteq\|B\|^M \text{ then } f(w,\|A\|^M)\subseteq f(w,\|A\wedge B\|^M), \tag{2}$$

Since $f(w,\|A\|^M)\neq\varnothing$, then if $f(w,\|A\|^M)\subseteq\|B\|^M$ then $f(w,\|A\|^M)\not\subseteq\neg\|B\|^M$,

or consequently $f(w,\|A\|^M)\not\subseteq\|\neg B\|^M$.

But by the condition (iii) on f,

$$\text{if } f(w,\|A\|^M)\not\subseteq\|\neg B\|^M \text{ then } f(w,\|A\wedge B\|^M)\subseteq f(w,\|A\|^M). \tag{3}$$

(2) and (3) together give us that:

$$\text{if } f(w,\|A\|^M)\subseteq\|B\|^M \text{ then } f(w,\|A\|^M)=f(w,\|A\wedge B\|^M). \tag{4}$$

Similarly,

$$\text{if } f(w,\|A\|^M)\subseteq\|\neg B\|^M \text{ then } f(w,\|A\|^M)=f(w,\|A\wedge\neg B\|^M). \tag{5}$$

By condition (iii) of f,

$$\text{if } f(w,\|A\|^M)\not\subseteq\|B\|^M \text{ then } f(w,\|A\wedge\neg B\|^M)\subseteq f(w,\|A\|^M).$$

So, if $f(w,\|A\|^M)\not\subseteq\|B\|^M$ and $f(w,\|A\|^M)\not\subseteq\|\neg B\|^M$ then

$$f(w,\|A\wedge\neg B\|^M)\subseteq f(w,\|A\|^M) \text{ and } f(w,\|A\wedge B\|^M)\subseteq f(w,\|A\|^M)$$

or consequently

$$f(w,\|A\wedge\neg B\|^M)\cup f(w,\|A\wedge B\|^M)\subseteq f(w,\|A\|^M). \tag{6}$$

From condition (iv) on f, $f(w,\|C\vee D\|^M)\subseteq f(w,\|C\|^M)\cup f(w,\|D\|^M)$.

Substituting $A \wedge B$ for C and $A \wedge \neg B$ for D, we get:

$$f(w, \|A\|^M) \subseteq f(w, \|A \wedge B\|^M) \cup f(w, \|A \wedge \neg B\|^M). \tag{7}$$

Combining (6) and (7), we get:

if $f(w, \|A\|^M) \not\subseteq \|B\|^M$ and $f(w, \|A\|^M) \not\subseteq \|\neg B\|^M$

then $f(w, \|A\|^M) = f(w, \|A \wedge B\|^M) \cup f(w, \|A \wedge \neg B\|^M)$. (8)

But given (1), the sentences (4), (5) and (8) imply that

$f(w, \|A\|^M) = f(w, \|A \wedge B\|^M)$ or $f(w, \|A\|^M) = f(w, \|A \wedge \neg B\|^M)$ or
$f(w, \|A\|^M) = f(w, \|A \wedge B\|^M) \cup f(w, \|A \wedge \neg B\|^M)$.

Definition 6: For specific $w \in W$, the binary relation E_f is given by

(i) If $w_1 \in f(w, \|A\|^M)$ then $E_f w w_1$

(ii) If $\vdash_w^M \neg(A \Rightarrow \neg B)$ then if $w_1 \in f(w, \|A\|^M)$ and $w_2 \in f(w, \|B\|^M)$ then $E_f w_1 w_2$

Theorem 6: If $M = \langle W, f, P \rangle$ is a $NP+$-model with a distinguished world w, then E_f is a reflexive, forward connected relation.

Proof:

Proof of reflexivity is trivial using the contrapositive of condition (v) and the fact that for primitive proposition p we must have $\vdash_w^M p$ or $\vdash_w^M \neg p$.

For forward connectedness we must show that if $E_f w w_1$ and $E_f w w_2$ then either $E_f w_1 w_2$ or $E_f w_2 w_1$.

Assume that $w_1 \in f(w, \|A\|^M)$ for some A and $w_2 \in f(w, \|B\|^M)$ for some B.

If either $\vdash_w^M \neg(A \Rightarrow \neg B)$ or $\vdash_w^M \neg(B \Rightarrow \neg A)$ then by condition (ii) in the definition of E_f, either $E_f w_1 w_2$ or $E_f w_2 w_1$, as required.

So assume not $\vdash_w^M \neg(A \Rightarrow \neg B)$ and not $\vdash_w^M \neg(B \Rightarrow \neg A)$.

Hence $f(w, \|A\|^M) \not\subseteq \|\neg B\|$ and by condition (iii) on f, $f(w, \|A \wedge \neg B\|^M) \subseteq f(w, \|A\|^M)$.

Similarly for not $\vdash_w^M B \Rightarrow A$ we get $f(w, \|\neg A \wedge B\|^M) \subseteq f(w, \|B\|^M)$.

By the theorem 2 of this section we get

(i) $f(w, \|A \vee B\|^M) = f(w, \|A\|^M)$ or

(ii) $f(w, \|A \vee B\|^M) = f(w, \|B\|^M)$ or

(iii) $f(w, \|A \vee B\|^M) = f(w, \|A\|^M) \cup f(w, \|B\|^M)$.

For case (i):

Since $w_1 \in f(w, \|A\|^M)$ (by the original assumption), we have that $w_1 \in f(w, \|A \vee B\|^M)$.

Also (by assumption) $w_2 \in f(w, \|B\|^M)$.

Since $\vdash_w^M B \Rightarrow A \vee B$, and hence $\vdash_w^M \neg(B \Rightarrow \neg(A \vee B))$, so by condition (ii) $E_f w_2 w_1$.

Case (ii) is proven analogously, yielding $E_f w_1 w_2$. Case (iii) is a proof by cases where the alternatives reduce to case (i) or case (ii).

With E_f defined as in the previous theorem, it follows easily that, for designated $w \in W$, if $E_f w w_1$ and $E_f w_1 w_2$ then it must be the case that $E_f w w_2$. Thus from the point of view of w, E_f is a reflexive, transitive and forward connected relation. This, informally, is why at w we can reason about formulae where no occurrence of a variable conditional falls within the scope of another. Furthermore, since we are interested only in the truth values of formulae at some distinguished world, the formulation of a model in terms of the function f is perfectly adequate for our purposes. We obtain:

Theorem 7: If $M = \langle W, f, P \rangle$ is an NP+-model then there is a E-model $M_E = \langle W, E_f, P \rangle$ such that for designated $w \in W$, M_f and M_E verify exactly the same sentences at w. That is, for $A \in L_P$, $\vdash_w^M A$ iff $\vdash_w^{M_E} A$.

Proof: The proof is straightforward, given that we do not have iterated modalities in L_P. For w_1, distinct from w, one could specify arbitrary values for worlds accessible from w_1. See theorem 5 for the general approach.

This result then is quite a bit weaker than that given for the opposite direction in theorem 5. Note that in going from an NP+-model to (supposedly) an E-model, we do not have transitivity of E_f in general. As a counterexample consider a NP+-model where $w, w_1, w_2 \in W$ are distinct, $f(w, \|A\|^M)$ and $f(w_1, \|A\|^M)$ are distinct and $w_1 \in f(w, \|A\|^M)$ and $w_2 \in f(w_1, \|A\|^M)$. Hence $E_f w w_1$ and $E_f w_1 w_2$ but we do not have $E_f w w_2$. In order to be able to fully obtain the relation E_f from f, and consequently be able to reason with nested variable conditionals, we require an additional condition on f:

Theorem 8: If $M = \langle W, f, P \rangle$ is a NP+-model and we define E_f as before, together with the condition:

(iii) If $w_1 \in f(w, \|A\|^M)$ and $w_2 \in f(w_1, \|B\|^M)$ then $w_2 \in f(w, \|B\|^M)$

then E_f is a transitive relation.

Proof: If $E_f w w_1$ then for some A, $w_1 \in f(w, \|A\|^M)$.

If $E_f w_1 w_2$ then for some B, $w_2 \in f(w_1, \|B\|^M)$.

By condition (iii) we obtain that $w_2 \in f(w, \|B\|^M)$.

By condition (i) we obtain that $E_f w w_2$.

Note that condition (iii) can be expressed equivalently as:

(iii') If $w_1 \in f(w, \|A\|^M)$ then $f(w_1, \|B\|^M) \subseteq f(w, \|B\|^M)$

or, assuming that we can order sets of worlds (using $\leq$ for notation – again see section 5) as:

(iii'') If $w_1 \leq w$ then $f(w_1, \|B\|^M) \subseteq f(w, \|B\|^M)$.

What this means is that as we consider less exceptional worlds, that $f(w, \|A\|^M)$ may decrease in size – this is shown explicitly in condition (iii'').[2] Now, a reasonable constraint would be that $f(w, \|A\|^M)$ not decrease to $\emptyset$. This may be stated as:

If $w_1 \in f(w, \|A\|^M)$ then $f(w_1, \|A\|^M) \neq \emptyset$

However if $w_1 \in f(w, \|A\|^M)$ then $\vdash^M_{w_1} A$, and from condition (v) on f we obtain that $f(w_1, \|A\|^M) \neq \emptyset$.

Given our interpretation of the conditional operator, it is difficult to see what use could be made of iterated occurrences. Also, as will become apparent, the ability to represent the variable conditional as holding in other possible states of affairs, in fact adds nothing of substance to what we can already say.

Lastly, bringing the two semantic approaches into line, we obtain:

Theorem 9: If $M = \langle W, f, P \rangle$ is an $NP+$-model then there is a E-model $M_E = \langle W, E_f, P \rangle$ such that M_f and M_E verify exactly the same sentences. That is, for all $w \in W$ and $A \in L_P$, $\vdash^M_w A$ iff $\vdash^{M_E}_w A$.

Proof: If E_f is defined as in definition 6, with condition (iii) as given in theorem 8, for all $w \in W$. The result follows easily.

2.3. The Conditional Logics NP and NP+

For this section, I first deal with the logic NP and only after soundness and completeness results have been obtained is NP+ considered. Since the proofs were originally with respect to NP, and the modifications to NP+ are straightforward, this permits minimal alteration of the original proofs.

Definition 1: The conditional logic *NP* is the smallest logic containing the standard propositional calculus (PC), and closed under the following axiom schemata and rule of inference.

ID	$A \Rightarrow A$
CC	$(A \Rightarrow B \wedge A \Rightarrow C) \supset (A \Rightarrow B \wedge C)$
RT	$A \Rightarrow B \supset (A \wedge B \Rightarrow C \supset A \Rightarrow C)$
CV	$\neg(A \Rightarrow B) \supset (A \Rightarrow C \supset A \wedge \neg B \Rightarrow C)$
CC′	$(A \Rightarrow C \wedge B \Rightarrow C) \supset (A \vee B \Rightarrow C)$
RCM	From $B \supset C$ infer $A \Rightarrow B \supset A \Rightarrow C$

(For naming the axioms and rules of inference, I have followed, and will follow, the conventions of [Chellas 75] and [Nute 80].)

The logic is, in the terminology of [Chellas 75], *normal* – that is, it is closed under the rules:

---

[2] I write $w_1 \leq w$ rather than $E_f w w_1$, because this ordering of worlds may be expressed without recourse to a (derived) accessibility relation.

RCEA From $A \equiv B$ infer $(A \Rightarrow C) \equiv (B \Rightarrow C)$.

RCK From $(B_1 \wedge \cdots \wedge B_n) \supset B$ infer $((A \Rightarrow B_1) \wedge \cdots \wedge (A \Rightarrow B_n)) \supset (A \Rightarrow B)$ for every $n \geq 0$.

From this it follows that equivalent propositions can be substituted into the antecedent or consequent of a variable conditional, and from this an easy inductive argument shows we have full substitution of equivalents.

A wff $A \in L_P$ is a *theorem* of *NP*, just when $A \in NP$. We write $\vdash_{NP} A$ or just $\vdash A$ when no confusion arises. A is *derivable* in *NP* from a set of wffs Γ iff there are $A_1, \ldots, A_n \in \Gamma$ such that $\vdash_{NP} (A_1 \wedge \cdots \wedge A_n) \supset A$. A set of wffs is *consistent* in *NP* in the case that not every wff is derivable from it. A set Γ of wffs of L_P is *maximal NP-consistent*[3] iff Γ is consistent and for every $A \in L_P$, either $A \in \Gamma$ or $\neg A \in \Gamma$. $|A|_{NP}$ will be used to stand for the class of maximal *NP*-consistent sets of wffs containing A. Finally, a set of wffs Δ is a *maximal (consistent) extension* of a consistent set of wffs Γ if Δ is maximal consistent, and $\Gamma \subseteq \Delta$.

From Lindenbaum's lemma, we have that if Γ is a (*NP*-)consistent set of wffs, then there is a maximal consistent set of wffs Δ such that $\Gamma \subseteq \Delta$. In addition, A is derivable from Γ iff A belongs to every maximal extension of Γ. Hence $\vdash A$ iff A belongs to every maximal consistent set of sentences.

Theorem 1: Every member of *NP* is (*NP*-)valid.

Proof: Every tautology of PC is clearly valid. The axioms of *NP* are easily shown to be valid, and the rules MP and RCM are easily shown to preserve validity. Hence by the definition of *NP* every member of *NP* is valid.

Completeness of *NP* is proven by showing that for any consistent set of formulae Γ there is a model, called the *canonical model*, in which there is a world in which every member of Γ is true. This approach is standard for propositional conditional logics (see, for example, [Chellas 75], [Lewis 73], [Nute 80]), and indeed for modal logics in general.

Definition 2: The *canonical model* $M = \langle W, f, P \rangle$ for *NP* is defined as follows:

(i) $W = \{w \mid w \text{ is a maximal consistent set of wffs}\}$,

(ii) $f(w, |A|_{NP}) = \{w_1 \in W \mid \{B \mid A \Rightarrow B \in w\} \subseteq w_1\}$,

(iii) $P(p_i) = \{w \in W \mid p_i \in w\}$.

Lemma 1: If Γ is a consistent set of wffs of L_P that contains the wff $\neg(A \Rightarrow B)$, then $\{C \mid A \Rightarrow C \in \Gamma\} \cup \{\neg B\}$ is consistent.

Proof: Let $\Lambda = \{C \mid A \Rightarrow C \in \Gamma\}$ and assume that $\Lambda \cup \{\neg B\}$ is not consistent.

So there are $B_1, \ldots, B_n$ of Λ such that $\vdash \neg(B_1 \wedge \cdots \wedge B_n \wedge \neg B)$ by the compactness theorem.

---

[3] For notational convenience I will drop the prefix *NP* when there is no confusion. A similar remark applies when we consider the first-order logic *N*.

So, by PC, $\vdash (B_1 \wedge \cdots \wedge B_n) \supset B$

and by RCK, $\vdash (A \Rightarrow B_1 \wedge \cdots \wedge A \Rightarrow B_n) \supset A \Rightarrow B$.

Hence $\vdash \neg(A \Rightarrow B_1 \wedge \cdots \wedge A \Rightarrow B_n \wedge \neg(A \Rightarrow B))$,

and so $\{A \Rightarrow B_1, \ldots, A \Rightarrow B_n, \neg(A \Rightarrow B)\}$ is not consistent.

Since this set is a subset of Γ, Γ is inconsistent.

But this proves the contrapositive of what we set out to prove, and hence the lemma.

Theorem 2: Let $M = \langle W, f, P \rangle$ be the canonical model for *NP*. Then for $A \Rightarrow B \in L_p$ and $w \in W$, $f(w, |A|_{NP}) \subseteq |B|_{NP}$ iff $A \Rightarrow B \in w$.

Proof: Assume that $A \Rightarrow B \in w$.

If $w_2 \in f(w, |A|_{NP})$ then $w_2 \in \{w_1 \in W \mid \{C \mid A \Rightarrow C \in w\} \subseteq w_1\}$ by the definition of the canonical model.

So $\{C \mid A \Rightarrow C \in w\} \subseteq w_2$.

In particular $B \in w_2$, and so $w_2 \in |B|_{NP}$.

And so $f(w, |A|_{NP}) \subseteq |B|_{NP}$.

Conversely suppose that $A \Rightarrow B \notin w$.

Then $\neg(A \Rightarrow B) \in w$ since w is maximal consistent.

So $\{C \mid A \Rightarrow C \in w\} \cup \{\neg B\}$ is consistent by lemma 1, and

$\{C \mid A \Rightarrow C \in w\} \cup \{\neg B\}$ has a maximal consistent extension w_1 by Lindenbaum's lemma.

Since $\{C \mid A \Rightarrow C \in w\} \subseteq w_1$, $w_1 \in f(w, |A|_{NP})$. (1)

Since $\neg B \in w_1$, $B \notin w_1$, and $w_1 \notin |B|_{NP}$. (2)

But (1) and (2) together imply $f(w, |A|_{NP}) \not\subseteq |B|_{NP}$.

Theorem 3: Let $M = \langle W, f, P \rangle$ be the canonical model for *NP*. Then for every $A \in L_p$ and $w \in W$, $\vdash_w^M A$ iff $A \in w$.

Proof: The proof is the same as that of Theorem 2 in [Chellas 75], which gives the same result for a class of conditional logics. The proof is by induction; the only case of interest involves the conditional operator: $\vdash_w^M A \Rightarrow B$ iff $f(w, \|A\|^M) \subseteq \|B\|^M$ iff $f(w, |A|_{NP}) \subseteq |B|_{NP}$ (by the inductive hypothesis) iff $A \Rightarrow B \in w$ (by the previous theorem).

The theorem implies that $\|A\|^M = |A|_{NP}$ for every wff $A \in L_p$. We also obtain:

Corollary: Any wff A is valid in the canonical model for *NP* iff $\vdash_{NP} A$.

NP is complete if every valid wff is a theorem of *NP*, or, conversely, if every non-theorem of *NP* is invalid in some model. By the above corollary, any non-theorem is invalid in the canonical model. Hence it remains only to be shown that the canonical model for *NP* is in fact an *NP*-model.

Theorem 4: The canonical model for *NP* is a *NP*-model.

Proof: The canonical model for *NP* is clearly a model. We need to show that conditions (i)-(iv) of a *NP*-model are met:

(i) Let $w_1 \in f(w, |A|_{NP})$.
Then $\{C \mid A \Rightarrow C \in w\} \subseteq w_1$ by the definition of the canonical model.
But $A \Rightarrow A \in w$, by ID and the fact that w is maximal consistent.
So $A \in w_1$, and hence $w_1 \in |A|_{NP}$ and $f(w, |A|_{NP}) \subseteq |A|_{NP}$.

(ii) Assume that $f(w, |A|_{NP}) \subseteq |B|_{NP}$.
We want to show that $f(w, |A|_{NP}) \subseteq f(w, |A \wedge B|_{NP})$.
Equivalently, if we assume that $w_1 \in f(w, |A|_{NP})$ we must show that $w_1 \in f(w, |A \wedge B|_{NP})$.
That is, we need to show that $\{C \mid A \wedge B \Rightarrow C \in w\} \subseteq w_1$.
Consider $C' \in \{C \mid A \wedge B \Rightarrow C \in w\}$.
Hence $A \wedge B \Rightarrow C' \in w$.
Since also $A \Rightarrow B \in w$, by RT and MP we get that $A \Rightarrow C' \in e$, and so $f(w, |A|_{NP}) \subseteq |C'|_{NP}$.
But by the initial assumption $w_1 \in f(w, |A|_{NP})$ and hence $w_1 \in |C'|_{NP}$, and so $C' \in w_1$.
Since $C' \in \{C \mid A \wedge B \Rightarrow C \in w\}$ yields $C' \in w_1$, we get that $\{C \mid A \wedge B \Rightarrow C \in w\} \subseteq w_1$, which was what we wanted to show.

(iii) Assume $f(w, |A|_{NP}) \not\subseteq |B|_{NP}$.
Thus $A \Rightarrow B \notin w$.
Since w is maximal consistent, $\neg(A \Rightarrow B) \in w$.
If also $f(w, |A|_{NP}) \subseteq |C|_{NP}$ then $A \Rightarrow C \in w$.
Then, by CV and MP, $A \wedge \neg B \Rightarrow C \in w$ and $f(w, |A \wedge \neg B|_{NP}) \subseteq |C|_{NP}$.
So $f(w, |A \wedge \neg B|_{NP}) \subseteq f(w, |A|_{NP})$.

(iv) Let $f(w, |A|_{NP}) \subseteq |C|_{NP}$ and $f(w, |B|_{NP}) \subseteq |C|_{NP}$.
Then $A \Rightarrow C \in w$, $B \Rightarrow C \in w$.
By CC′ and MP, $A \vee B \Rightarrow C \in w$, and hence $f(w, |A \vee B|_{NP}) \subseteq |C|_{NP}$,
and so $f(w, |A \vee B|_{NP}) \subseteq f(w, |A|_{NP}) \cup f(w, |B|_{NP})$.

Definition 3: The conditional logic *NP+* is the smallest logic containing *NP* and closed under the additional axiom schemata

REF $\quad (\neg A \Rightarrow A) \supset A$

CP $\quad A \Rightarrow C \supset (\neg(A \vee B \Rightarrow C) \supset \neg C \Rightarrow \neg A)$

REF corresponds to an axiom of reflexivity. It was omitted from the original systems for reasons of economy: it didn't seem to be needed, and seemed to slightly complicate the systems, and so was not included. However, with the development of default reasoning in the systems, it turned out that things were made a bit more complex with the omission of the axiom (see [Delgrande 88]); and so it has here

been included. CP corresponds to a restricted form of the contrapositive. It is required only in a very obscure set of transitive denied defaults (see theorem 13, section 2.5); unfortunately certain equivalence proofs in recent (area 4 in section 1) work, not covered here, relies on this result.

Definitions for derivability, theoremhood, etc., will of course be exactly as before, except subscripted (or whatever) with $NP+$. The soundness result (with respect to the set of $NP+$-models) is proven by the usual inductive argument:

Theorem 5: Every member of $NP+$ is valid with respect to the set of $NP+$-models.

For completeness the only step we will have to take, in addition to the completeness proof of NP, is to show that the canonical model for $NP+$ is in fact a $NP+$-model. (See [Hughes and Cresswell 84] for a full development of this approach.)

Theorem 6: $NP+$ is complete with respect to the class of $NP+$-models.

Proof: We need to prove that in the canonical model for $NP+$ that the two additional conditions on f are met.

(v) We have that $A \supset \neg(A \Rightarrow \neg A) \in w$ for every $w \in W$.
Assume that $A \in w$; hence $\neg(A \Rightarrow \neg A) \in w$.
Thus $\{C \mid A \Rightarrow C \in w\} \cup \{A\}$ is consistent by lemma 1.
So $\{C \mid A \Rightarrow C \in w\} \cup \{A\}$ has a maximal consistent extension w_2.
But by definition $f(w, |A|_{NP}) = \{w_1 \in W \mid \{B \mid A \Rightarrow B \in w\} \subseteq w_1\}$,
and $w_2 \in \{w_1 \in W \mid \{B \mid A \Rightarrow B \in w\} \subseteq w_1\}$.
Hence $\{w_1 \in W \mid \{B \mid A \Rightarrow B \in w\} \subseteq w_1\} \neq \emptyset$.

(vi) Assume that $f(w, |A|_{NP}) \subseteq |C|_{NP}$ and $f(w, |A \vee B|_{NP}) \not\subseteq |C|_{NP}$.
Hence $A \Rightarrow C \in w$ and $A \vee B \Rightarrow C \notin w$.
Since w is maximal consistent, $\neg(A \vee B \Rightarrow C) \in w$.
Since $A \Rightarrow C \supset (\neg(A \vee B \Rightarrow C) \supset \neg C \Rightarrow \neg A) \in w$, we have that $\neg C \Rightarrow \neg A \in w$.
Assume that $w_1 \in f(w, |\neg C|_{NP})$.
So $\{P \mid \neg C \Rightarrow P \in w\} \subseteq w_1$.
But $\neg C \Rightarrow \neg A \in w$, and so $\neg A \in w_1$ and hence $w_1 \in |\neg A|_{NP}$.

By a standard result concerning a type of derived model called a *filtration* [Hughes and Cresswell 84], [Chellas 75] we obtain:

Theorem 9: NP and $NP+$ are decidable.

2.4. Some Results

All results is this section are proved using only the axioms for NP. First, the conditional operator can be tied to the more familiar modal notions of necessity or truth in all alternative states of affairs, and possibility in the following manner [Lewis 73]:

Definition:

$\Box A$ is defined as $\neg A \Rightarrow A$.

$\Diamond A$ is defined as $\neg(A \Rightarrow \neg A)$.

That is, if $\neg A \Rightarrow A$ is true at w then $f(w, \|\neg A\|^M) \subseteq \|A\|^M$, hence $\|\neg A\| = \emptyset$, hence $\|A\|^M = W$.

The following (unsurprising) results can be shown:

(i) If $\vdash A$ then $\vdash \Box A$.

(ii) $\vdash \Box(A \supset B) \supset (\Box A \supset \Box B)$.

As discussed, the world selection function f, as specified, is too weak to allow iterated modalities. If we were to allow iterated modalities, we would also obtain the somewhat curious result that

$$\vdash (A \Rightarrow B) \equiv \Box(A \Rightarrow B).$$

Thus, insofar as the variable conditional is concerned, augmenting the language to allow necessity adds virtually nothing with respect to expressibility.

Theorem 1:

RI	If $\vdash A \supset B$ then $\vdash A \Rightarrow B$
Weak CEM	$\vdash \Diamond A \supset (A \Rightarrow B \supset \neg(A \Rightarrow \neg B))$
CSO	$\vdash (A \Rightarrow B \wedge B \Rightarrow A) \supset (A \Rightarrow C \equiv B \Rightarrow C)$
RCEA	If $\vdash A \equiv B$ then $\vdash A \Rightarrow C \equiv B \Rightarrow C$
RCK	If $\vdash (B_1 \wedge \cdots \wedge B_n) \supset B$ then $\vdash (A \Rightarrow B_1 \wedge \cdots \wedge A \Rightarrow B_n) \supset A \Rightarrow B$
CI	$\vdash (A \Rightarrow (B \supset C)) \supset (A \Rightarrow B \supset A \Rightarrow C)$

Proof:

RI:

If $\vdash A \supset B$ then, via RCM, $\vdash A \Rightarrow A \supset A \Rightarrow B$.

ID gives $\vdash A \Rightarrow A$, and so by MP we get $\vdash A \Rightarrow B$.

Weak CEM:

1. $(B \wedge \neg B) \supset \neg A$	PC
2. $A \Rightarrow (B \wedge \neg B) \supset A \Rightarrow \neg A$	1,RCM

3. $((A \Rightarrow B) \wedge (A \Rightarrow \neg B)) \supset (A \Rightarrow (B \wedge \neg B))$ CC

4. $(A \Rightarrow B \wedge A \Rightarrow \neg B) \supset (A \Rightarrow \neg A)$ 3,2,PC

5. $(A \Rightarrow B) \supset ((A \Rightarrow \neg B) \supset (A \Rightarrow \neg A))$ 4,PC

6. $\neg(A \Rightarrow \neg A) \supset (A \Rightarrow B \supset \neg(A \Rightarrow \neg B))$ 5,PC

7. $\Diamond A \supset (A \Rightarrow B \supset \neg(A \Rightarrow \neg B))$ 6,Defn. poss.

CSO:

1. $\Diamond A \supset ((A \Rightarrow B) \supset \neg(A \Rightarrow \neg B))$ Weak CEM

2. $\neg(A \Rightarrow \neg B) \supset (A \Rightarrow C \supset A \wedge B \Rightarrow C)$ CV

3. $\Diamond A \supset (A \Rightarrow B \supset (A \Rightarrow C \supset A \wedge B \Rightarrow C))$ 1,2,PC

4. $A \Rightarrow B \supset (A \wedge B \Rightarrow C \supset A \Rightarrow C)$ RT

5. $\Diamond A \supset (A \Rightarrow B \supset (A \wedge B \Rightarrow C \equiv A \Rightarrow C))$ 3,4,PC

6. $\Diamond B \supset (B \Rightarrow A \supset (A \wedge B \Rightarrow C \equiv B \Rightarrow C))$ 5,Subst. vars.

7. $(\Diamond A \wedge \Diamond B) \supset ((A \Rightarrow B \wedge B \Rightarrow A) \supset (A \Rightarrow C \equiv B \Rightarrow C))$ 5,6,PC

8. $\neg\Diamond A \supset ((A \wedge \neg A \Rightarrow C) \supset A \Rightarrow C)$ RT

9. $A \wedge \neg A \Rightarrow C$ RI

10. $\neg\Diamond A \supset A \Rightarrow C$ 8,9,PC

11. $\neg\Diamond B \supset B \Rightarrow C$ 10,Subst. vars.

12. $(\neg\Diamond A \wedge \neg\Diamond B) \supset (A \Rightarrow C \equiv B \Rightarrow C)$ 10,11,PC

13. $(A \Rightarrow B \wedge B \Rightarrow A) \supset (A \Rightarrow C \equiv B \Rightarrow C)$ 7,12,PC

RCEA:

1. $\vdash A \equiv B$ Supposition

2. $\vdash A \Rightarrow B \wedge B \Rightarrow A$ 1,RI

3. $\vdash (A \Rightarrow B \wedge B \Rightarrow A) \supset (A \Rightarrow C \equiv B \Rightarrow C)$ CSO

4. $\vdash A \Rightarrow C \equiv B \Rightarrow C$ 3,2,MP

RCK:

Proof is by induction on n. For $n = 1$ the result is RCM. For $n = k+1$:

1. $\vdash (B_1 \wedge \cdots \wedge B_{k+1}) \supset B$ Supposition

2. $\vdash ((A \Rightarrow B_1) \wedge \cdots \wedge (A \Rightarrow (B_k \wedge B_{k+1}))) \supset A \Rightarrow B$ — 1,Ind. hyp.

3. $\vdash ((A \Rightarrow B_k) \wedge (A \Rightarrow B_{k+1})) \supset (A \Rightarrow B_k \wedge B_{k+1})$ — CC

4. $\vdash ((A \Rightarrow B_1) \wedge \cdots \wedge (A \Rightarrow B_k) \wedge (A \Rightarrow B_{k+1})) \supset A \Rightarrow B$ — 2,3,PC

CI:

1. $(B \wedge C) \supset C$ — PC

2. $A \Rightarrow (B \wedge C) \supset A \Rightarrow C$ — 1,RCM

3. $((A \Rightarrow (\neg B \vee C)) \wedge A \Rightarrow B) \supset (A \Rightarrow ((\neg B \vee C) \wedge B))$ — CC

4. $((A \Rightarrow (\neg B \vee C)) \wedge A \Rightarrow B) \supset (A \Rightarrow B \wedge C)$ — 3,Subst. equiv.

5. $(A \Rightarrow (\neg B \vee C) \wedge A \Rightarrow B) \supset (A \Rightarrow C)$ — 4,2,PC

6. $(A \Rightarrow (\neg B \vee C)) \supset ((A \Rightarrow B) \supset (A \Rightarrow C))$ — 5,PC

7. $(A \Rightarrow (B \supset C)) \supset ((A \Rightarrow B) \supset (A \Rightarrow C))$ — 6,PC

Theorem 2:

(i) If $\vdash A$ then $\vdash \Box A$

(ii) $\vdash A \Rightarrow B$ iff $\vdash A \supset B$.

(iii) $\Box A \supset \neg A \Rightarrow B$

(iv) $\Box(A \supset B) \supset A \Rightarrow B$

(v) $\Box A \supset B \Rightarrow A$

(vi) $\Box A \supset \Box(B \supset A)$

(vii) $\Box(A \supset B) \supset (\Box A \supset \Box B)$

Proof:

(i):

Assume that $\vdash A$.

Hence by PC we get $\vdash \neg A \supset A$, and by RI we get $\vdash \neg A \Rightarrow A$.

From the definition of $\Box$ we get $\vdash \Box A$.

(ii)

If $\vdash A \supset B$ then $\vdash A \Rightarrow B$ by RI.

For the other direction, observe that if $\supset$ is substituted for $\Rightarrow$ in each of the axioms of N, that the resulting formulae are all valid in PC. Moreover RCM with $\Rightarrow$ replaced with $\supset$ is a valid rule of inference in PC. The desired result follows easily by induction.

(iii):

1. $(\neg A \Rightarrow A) \supset ((\neg A \wedge A \Rightarrow B) \supset \neg A \Rightarrow B)$	RT
2. $\neg A \wedge A \Rightarrow B$	PC,RI
3. $(\neg A \Rightarrow A) \supset (\neg A \Rightarrow B)$	1,2,PC

(iv):

1. $\Box(A \Rightarrow B) \equiv ((A \wedge \neg B) \Rightarrow (A \supset B))$	Defn. nec., Subst. equiv.
2. $\Box(A \Rightarrow B) \equiv ((A \wedge \neg B \Rightarrow A) \supset (A \wedge \neg B \Rightarrow B))$	1,CI
3. $A \wedge \neg B \Rightarrow A$	PC,RI
4. $\Box(A \Rightarrow B) \equiv (A \wedge \neg B \Rightarrow B)$	2,3,PC
5. $((A \wedge \neg B \Rightarrow B) \wedge (A \wedge B \Rightarrow B)) \supset (A \Rightarrow B)$	CC
6. $A \wedge B \Rightarrow B$	PC,RI
7. $A \wedge \neg B \Rightarrow B \supset A \Rightarrow B$	5,6,PC
8. $\Box(A \supset B) \supset A \Rightarrow B$	4,7,PC

(v):

1. $\Box A \supset (\neg A \Rightarrow (B \wedge \neg A))$	(iii) above
2. $(B \wedge \neg A) \Rightarrow \neg A$	PC,RI
3. $\Box A \supset ((\neg A \Rightarrow B \wedge \neg A) \vee (\neg A \wedge B \Rightarrow \neg A))$	1,2,PC
4. $((\neg A \Rightarrow B \wedge A) \vee (\neg A \wedge B \Rightarrow \neg A)) \supset (\neg A \Rightarrow A \equiv \neg A \wedge B \Rightarrow A)$	CSO
5. $\Box A \supset (\neg A \Rightarrow A \equiv \neg A \wedge B \Rightarrow A)$	3,4,PC
6. $\Box A \supset \neg A \wedge B \Rightarrow A$	5,Defn. nec.,PC
7. $\neg A \wedge B \Rightarrow B$	PC,RI
8. $\Box A \supset ((\neg A \wedge B \Rightarrow A) \vee (\neg A \wedge B \Rightarrow B))$	6,7,PC
9. $\Box A \supset ((\neg A \wedge B) \Rightarrow (A \wedge B))$	8,CC,PC
10. $((\neg A \wedge B \Rightarrow A \wedge B) \vee (A \wedge B \Rightarrow A \wedge B)) \supset B \Rightarrow B \wedge A$	CC'
11. $(\neg A \wedge B \Rightarrow A \wedge B) \supset (B \Rightarrow B \wedge A)$	10,ID,PC
12. $\Box A \supset (B \Rightarrow B \wedge A)$	9,11,PC
13. $((B \wedge A \Rightarrow B) \vee (B \Rightarrow B \wedge A)) \supset (B \wedge A \Rightarrow A \equiv B \Rightarrow A)$	CSO

14. $B \wedge A \Rightarrow B$	PC,RI
15. $(B \Rightarrow B \wedge A) \supset (B \wedge A \Rightarrow A \equiv B \Rightarrow A)$	13,14,PC
16. $B \wedge A \Rightarrow A$	PC,RI
17. $(B \Rightarrow B \wedge A) \supset (B \Rightarrow A)$	15,16,PC
18. $\Box A \supset B \Rightarrow A$	12,17,PC

(vi):

1. $\neg A \Rightarrow A \supset (\neg(B \supset A) \Rightarrow A)$	(v) above
2. $A \supset (B \supset A)$	PC
3. $(\neg(B \supset A) \Rightarrow A) \supset (\neg(B \wedge A) \Rightarrow (B \supset A))$	2,RCM
4. $\neg A \Rightarrow A \supset (\neg(B \supset A) \Rightarrow (B \supset A))$	1,3,PC
5. $\Box A \supset \Box(B \supset A)$	Defn. nec.

(vii):

1. $\Box(A \supset B) \supset \Box(\neg B \supset \neg A)$	Defn. nec.,Subst. equiv.
2. $\Box(\neg A \supset \neg B) \supset \neg B \Rightarrow \neg A$	(iv) above
3. $\Box(A \supset B) \supset \neg B \Rightarrow \neg A$	1,2,PC
4. $\Box A \supset \neg A \Rightarrow \neg B$	(iii) above
5. $(\neg A \Rightarrow B \wedge \neg B \Rightarrow A) \supset (\neg A \Rightarrow A \equiv \neg B \Rightarrow A)$	CSO
6. $\Box(A \supset B) \supset (\Box A \supset (\neg A \Rightarrow A \equiv \neg B \Rightarrow A))$	3,4,5,PC
7. $\neg B \Rightarrow (A \supset B) \supset (\neg B \Rightarrow A \supset \neg B \Rightarrow B)$	CI
8. $\neg B \Rightarrow A \supset ((A \wedge \neg B \Rightarrow (A \supset B)) \supset (\neg B \Rightarrow (A \supset B)))$	RT
9. $\neg B \Rightarrow A \supset (\Box(A \supset B) \supset (\neg B \Rightarrow (A \supset B)))$	8,Defn. nec.
10. $\Box(A \supset B) \supset ((\neg B \Rightarrow A) \supset (\neg B \Rightarrow (A \supset B)))$	9,PC
11. $\Box(A \supset B) \supset ((\neg B \Rightarrow A) \supset (\neg B \Rightarrow A \supset \neg B \Rightarrow B))$	10,7,PC
12. $\Box(A \supset B) \supset (\neg B \Rightarrow A \supset \neg B \Rightarrow B)$	11,PC
13. $\Box(A \supset B) \supset (\Box A \supset (\Box A \supset \neg B \Rightarrow B))$	6,12,PC
14. $\Box(A \supset B) \supset (\Box A \supset \Box B)$	15,PC,Defn. nec.

2.5. Some More Results

This section describes some minor, but useful results. Most are concerned either with reasoning about the antecedent of a conditional or with transitive defaults (i.e. defaults of the form $A \Rightarrow B, B \Rightarrow C$ or $\neg(A \Rightarrow \neg B), \neg(B \Rightarrow \neg C)$). The first theorem is a useful intermediate result used in a number of the other theorems; in addition however, it provides a key point of linkage between the two approaches for default reasoning (section 3). Theorems 5, 6, the corollary to theorem 11, and theorems 12, 18 and 20 all deal with eliminating disjunctions in an antecedent of a default. Theorems 7 and 17 deal with introducing disjunctions in a default (as of course does axiom CC'). Theorem 7 deals with transitive defaults, while theorems 8, 9, 13, and 14 deal with denied transitivities in defaults. Theorems 10, 19, and 21 concern simplifying the form of defaults with disjunctions in the antecedent. Lastly note that only theorem 13 requires the axiom CP; all the others are theorems of the more limited system *NP*. As a point of interest, most of the transitivities were proven in an attempt to prove CP in the system *NP*.

In summary, these results are:

Theorem 1: $\vdash_N (A \wedge B) \Rightarrow C \supset A \Rightarrow (B \supset C)$.

Theorem 2: $\vdash_N A \Rightarrow (B \supset C) \supset (\neg(A \Rightarrow \neg B) \supset (A \wedge B) \Rightarrow C)$.

Theorem 3: $\vdash_N (\neg(A \Rightarrow (B \supset C)) \wedge B \Rightarrow C) \supset B \Rightarrow \neg A$

Theorem 4: $\vdash_N ((A \wedge B \Rightarrow C) \wedge \neg(B \Rightarrow C)) \supset (B \wedge \neg C) \Rightarrow \neg A$

Corollary: $\vdash_N (B \Rightarrow (A \supset C) \wedge \neg(B \Rightarrow C)) \supset (B \wedge \neg C) \Rightarrow \neg A$

Theorem 5: $\vdash_N (A \vee C) \Rightarrow C \supset (\neg(A \Rightarrow C) \supset C \Rightarrow \neg A)$

Theorem 6: $\vdash_N (A \vee C) \Rightarrow C \wedge \neg(C \Rightarrow \neg A) \supset (C \Rightarrow B \supset A \Rightarrow B)$

Theorem 7: $\vdash_N A \Rightarrow B \supset (B \Rightarrow C \supset (A \vee C \Leftrightarrow C))$

Theorem 8: $\vdash_N (A \Rightarrow B \wedge B \Rightarrow C \wedge A \Rightarrow \neg C) \supset B \Rightarrow \neg A$

Corollary: $\vdash_N (\neg(A \Rightarrow \neg B) \wedge B \Rightarrow C \wedge A \Rightarrow \neg C) \supset B \Rightarrow \neg A$

Corollary: $\vdash_N (A \Rightarrow B \wedge B \Rightarrow C \wedge \neg(A \Rightarrow C)) \supset B \Rightarrow \neg A$

Theorem 9: $\vdash_N (\neg(A \Rightarrow \neg B) \wedge \neg(B \Rightarrow \neg C) \wedge A \Rightarrow \neg C) \supset \neg(B \Rightarrow A)$

Theorem 10: $\vdash_N (A \vee B \Rightarrow C \wedge \neg(A \Rightarrow C)) \supset A \vee B \Leftrightarrow B$

Theorem 11: $\vdash_N (A \vee B \Leftrightarrow A) \vee (A \vee B \Leftrightarrow B) \vee (A \vee B \Rightarrow C \equiv (A \Rightarrow C \wedge B \Rightarrow C))$

Corollary (ND): $\vdash_N A \vee B \Rightarrow C \supset (A \Rightarrow C \vee B \Rightarrow C)$

Theorem 12: $\vdash_N (A \vee B \Rightarrow C \wedge \neg(A \Rightarrow \neg B)) \supset B \Rightarrow C$

Corollary: $\vdash_N \neg(A \Rightarrow \neg B) \wedge \neg(B \Rightarrow C) \supset \neg(A \vee B \Rightarrow C)$

Theorem 13: $\vdash_{N+} (\neg(A \Rightarrow \neg B) \wedge \neg(B \Rightarrow \neg C) \wedge A \Rightarrow \neg C) \supset C \Rightarrow \neg A$

Theorem 14: $\vdash_N (A \Rightarrow B \wedge B \Rightarrow C \wedge \neg(A \Rightarrow C)) \supset C \Rightarrow \neg A$

Theorem 15: $\vdash_N A \Rightarrow C \supset (A \wedge B \Rightarrow C \vee A \wedge \neg B \Rightarrow C)$

Theorem 16: $\vdash_N (A \wedge B \Rightarrow C \wedge A \wedge \neg B \Rightarrow C) \supset A \Rightarrow C$

Theorem 17: $\vdash_N (A \Rightarrow B \wedge B \Rightarrow C) \supset A \vee B \Rightarrow C$

Theorem 18: $\vdash_N A \vee B \Rightarrow C \supset (A \Rightarrow C \vee C \Rightarrow \neg A)$

Theorem 19: $\vdash_N A \vee B \Rightarrow C \supset A \vee C \Leftrightarrow C$

Theorem 20: $\vdash_N A \vee B \Rightarrow \neg B \supset A \Rightarrow \neg B$

Theorem 21: $\vdash_N A \vee B \Rightarrow \neg B \supset A \vee B \Rightarrow A$

In more detail we have:

Theorem 1: $(A \wedge B) \Rightarrow C \supset A \Rightarrow (B \supset C)$

Proof:

1. $(A \wedge B) \Rightarrow C$	Supp.
2. $(A \wedge B) \Rightarrow (\neg B \vee C)$	1,PC,RCM,MP
3. $(A \wedge \neg B) \Rightarrow \neg B$	PC,RCM,MP
4. $(A \wedge \neg B) \Rightarrow (\neg B \vee C)$	3,PC,RCM,MP
5. $A \Rightarrow (\neg B \vee C)$	2,4,CC',MP
6. $((A \wedge B) \Rightarrow C) \supset (A \Rightarrow (\neg B \vee C))$	1,5,Disch.
7. $((A \wedge B) \Rightarrow C) \supset (A \Rightarrow (B \supset C))$	6,PC,Subst. Equiv.

The next theorem provides a partial converse to the previous.

Theorem 2: $A \Rightarrow (B \supset C) \supset (\neg(A \Rightarrow \neg B) \supset (A \wedge B) \Rightarrow C)$

Proof:

1. $A \Rightarrow (B \supset C)$	Supp.
2. $\neg(A \Rightarrow \neg B)$	Supp.
3. $(A \wedge B) \Rightarrow (B \supset C)$	1,2,CV,MP
4. $(A \wedge B) \Rightarrow B \supset (A \wedge B) \Rightarrow C$	3,CI
5. $(A \wedge B) \Rightarrow B$	RI
6. $(A \wedge B) \Rightarrow C$	4,5,MP
7. $A \Rightarrow (B \supset C) \supset (\neg(A \Rightarrow \neg B) \supset (A \wedge B) \Rightarrow C$	1,2,6,Disch

Theorem 3: $(\neg(A \Rightarrow (B \supset C)) \wedge B \Rightarrow C) \supset B \Rightarrow \neg A$

Proof:

1. $\neg(B \Rightarrow \neg A) \supset (B \Rightarrow C \supset A \wedge B \Rightarrow C)$	CV
2. $(A \wedge B \Rightarrow C) \supset (A \Rightarrow (B \supset C))$	Thm. 1
3. $\neg(B \Rightarrow \neg A) \supset (B \Rightarrow C \supset A \Rightarrow (B \supset C))$	1,2,Trans
4. $(\neg(A \Rightarrow (B \supset C)) \wedge B \Rightarrow C) \supset B \Rightarrow \neg A$	3,PC

Theorem 4: $((A \wedge B \Rightarrow C) \wedge \neg(B \Rightarrow C)) \supset (B \wedge \neg C) \Rightarrow \neg A$

Proof:

1. $A \wedge B \Rightarrow C \supset B \Rightarrow (C \vee \neg A)$	Thm. 1
2. $(\neg(B \Rightarrow C) \supset B \Rightarrow (C \vee \neg A)) \supset ((B \wedge \neg C) \Rightarrow (C \vee \neg A))$	CV
3. $(\neg(B \Rightarrow C) \supset A \wedge B \Rightarrow C) \supset ((B \wedge \neg C) \Rightarrow (C \vee \neg A))$	1,2,PC
4. $((B \wedge \neg C) \Rightarrow (C \vee \neg A)) \equiv ((B \wedge \neg C) \Rightarrow (\neg C \supset \neg A))$	Subst Equiv
5. $(B \wedge \neg C) \Rightarrow (\neg C \supset \neg A) \supset (((B \wedge \neg C) \Rightarrow \neg C) \supset (B \wedge \neg C \Rightarrow \neg A))$	CI
6. $(B \wedge \neg C) \Rightarrow (C \vee \neg A) \supset (((B \wedge \neg C) \Rightarrow \neg C) \supset (B \wedge \neg C \Rightarrow \neg A))$	4,5,PC
7. $(\neg(B \Rightarrow C) \vee (A \wedge B) \Rightarrow C) \supset (((B \wedge \neg C) \Rightarrow \neg C) \supset ((B \wedge \neg C) \Rightarrow \neg A))$	3,6,PC
8. $(\neg(B \Rightarrow C) \wedge (A \wedge B) \Rightarrow C) \supset ((B \wedge \neg C) \Rightarrow \neg A)$	7,R1,PC

Note that from line 1 of the proof we obtain the following result, which is reminiscent in form to theorem 3:

Corollary: $(B \Rightarrow (A \supset C) \wedge \neg(B \Rightarrow C)) \supset (B \wedge \neg C) \Rightarrow \neg A$

Theorem 5: $(A \vee C) \Rightarrow C \supset (\neg(A \Rightarrow C) \supset C \Rightarrow \neg A)$

Proof:

1. $\neg(A \vee C \Rightarrow \neg A) \supset (A \vee C \Rightarrow C \supset (((A \vee C) \wedge A) \Rightarrow C))$	CV
2. $\neg(A \vee C \Rightarrow \neg A) \supset (A \vee C \Rightarrow C \supset A \Rightarrow C)$	1,PC
3. $(A \vee C \Rightarrow C \wedge C \Rightarrow A \vee C) \supset (\neg(C \Rightarrow \neg A) \supset \neg(A \vee C \Rightarrow \neg A))$	CSO
4. $A \vee C \Rightarrow C \supset (\neg(C \Rightarrow \neg A) \supset \neg(A \vee C \Rightarrow \neg A))$	3,PC
5. $A \vee C \Rightarrow C \supset (\neg(C \Rightarrow \neg A) \supset (A \vee C \Rightarrow C \supset A \Rightarrow C))$	4,2,Trans.
6. $A \vee C \Rightarrow C \supset (\neg(A \Rightarrow C) \supset C \Rightarrow \neg A)$	5,PC

The next theorem is similar to the previous, in that it has the same antecedents. Basically it serves as an analogue to *CSO* (since $A \vee C \Rightarrow C$ is equivalent to $(A \vee C \Rightarrow C) \wedge (C \Rightarrow A \vee C)$) which allows disjunction to be eliminated from the antecedent of a conditional. Note that the preceding theorem is equivalent to $(A \vee C) \Rightarrow C \supset (\neg(C \Rightarrow \neg A) \supset (A \Rightarrow C))$.

Theorem 6: $(A \vee C) \Rightarrow C \wedge \neg(C \Rightarrow \neg A) \supset (C \Rightarrow B \supset A \Rightarrow B)$

Proof:

1. $A \vee C \Rightarrow C$	Supp.
2. $\neg(C \Rightarrow \neg A)$	Supp.
3. $A \vee C \Rightarrow C \supset (\neg(C \Rightarrow \neg A) \supset A \Rightarrow C)$	Thm 6,PC
4. $A \Rightarrow C$	1,2,3,MP
5. $C \Rightarrow B$	Supp.
6. $(C \Rightarrow B \wedge \neg(C \Rightarrow \neg A)) \supset A \wedge C \Rightarrow B$	CV
7. $A \wedge C \Rightarrow B$	5,2,6,MP
8. $((A \Rightarrow C) \wedge (A \wedge C \Rightarrow B)) \supset A \Rightarrow B$	RT
9. $A \Rightarrow B$	4,7,8,MP
10. $(A \vee C \Rightarrow C \wedge \neg(C \Rightarrow \neg A)) \supset (C \Rightarrow B \supset A \Rightarrow B)$	9,1,2,5,Disch

Theorem 7: $A \Rightarrow B \supset (B \Rightarrow C \supset (A \vee C \Leftrightarrow C))$

Proof:

1. $A \Rightarrow B$	Supp.
2. $B \Rightarrow C$	Supp.
3. $B \Rightarrow B$	ID
4. $(A \Rightarrow B \wedge B \Rightarrow B) \supset (A \vee B) \Rightarrow B$	CC'
5. $(A \vee B) \Rightarrow B$	1,3,4,MP
6. $B \Rightarrow (A \vee B)$	PC,RI
7. $(A \vee B) \Leftrightarrow B$	5,6,Defn <=>
8. $(A \vee B) \Rightarrow C$	7,2,CSO
9. $(A \vee B \vee C) \Rightarrow C$	8, see lines 3-7
10. $(A \vee B) \Rightarrow (A \vee C)$	8,RCM

11. $(A \vee C) \Rightarrow (A \vee C)$	PC,RI
12. $(A \vee B \vee C) \Rightarrow (A \vee C)$	10,11,CC'
13. $(A \vee B \vee C) \Leftrightarrow (A \vee C)$	13+PC,RI
14. $(A \vee C) \Rightarrow C$	CSO,13,9
15. $C \Rightarrow (A \vee C)$	PC,RI
16. $(A \vee C) \Leftrightarrow C$	14,15,Defn <=>

Theorems 8 and 9, and 13 and 14, all deal one way or another with denied transitivities. Note that the first two and theorem 14 are proven using axioms of N only; the third is proven in the augmented system. Note also that theorem 8 relies on the fact (for application of CEM) that $A \wedge B$ is in fact possible – a minor complication that for simplicity I have decided to ignore.

Theorem 8: $(A \Rightarrow B \wedge B \Rightarrow C \wedge A \Rightarrow \neg C) \supset B \Rightarrow \neg A$

Proof:

1. $A \Rightarrow B \wedge A \Rightarrow \neg C$	Supp.
2. $(A \Rightarrow B \wedge A \Rightarrow \neg C) \supset A \wedge B \Rightarrow \neg C$	CV+CEM
3. $A \wedge B \Rightarrow \neg C$	1,2,MP
4. $B \Rightarrow (A \supset \neg C)$	3,Thm. 1
5. $B \Rightarrow (C \supset \neg A)$	4,Subst Equiv
6. $B \Rightarrow (C \supset \neg A) \supset (B \Rightarrow C \supset B \Rightarrow \neg A)$	CI
7. $B \Rightarrow C \supset B \Rightarrow \neg A$	5,6,MP
8. $(A \Rightarrow B \wedge A \Rightarrow \neg C \wedge B \Rightarrow C) \supset B \Rightarrow \neg A$	1,7,Disch.

Corollary: $(\neg(A \Rightarrow \neg B) \wedge B \Rightarrow C \wedge A \Rightarrow \neg C) \supset B \Rightarrow \neg A$

Corollary: $(A \Rightarrow B \wedge B \Rightarrow C \wedge \neg(A \Rightarrow C)) \supset B \Rightarrow \neg A$

These results aren't in fact corollaries, but rather are obtained by easy modifications to the proof of theorem 8. (That is, one of the conditionals in the first line of the proof can be replaced by the denial of the default without affecting the remainder of the proof.) Note that theorem 8 is also an easy consequence of theorem 3. That is, since

$(A \Rightarrow B \wedge A \Rightarrow \neg C) \supset A \Rightarrow (B \wedge \neg C)$ and

$A \Rightarrow (B \wedge \neg C) \supset \neg(A \Rightarrow \neg(B \wedge \neg C))$ and so

$A \Rightarrow (B \wedge \neg C) \supset \neg(A \Rightarrow (B \supset C))$,

we get $\neg(A \Rightarrow (B \supset C))$ from two of the antecedent conditions of theorem 8. This together with $B \Rightarrow C$ provide the antecedent conditions for theorem 3, from which $B \Rightarrow \neg A$ may be derived.

Theorem 9: $(\neg(A \Rightarrow \neg B) \wedge \neg(B \Rightarrow \neg C) \wedge A \Rightarrow \neg C) \supset \neg(B \Rightarrow A)$

Proof:

1. $\neg(A \Rightarrow \neg B) \wedge A \Rightarrow \neg C$	Supp.
2. $(\neg(A \Rightarrow \neg B) \wedge A \Rightarrow \neg C) \supset A \wedge B \Rightarrow \neg C$	CV
3. $A \wedge B \Rightarrow \neg C$	1,2,MP
4. $B \Rightarrow (A \supset \neg C)$	4,Thm. 1
5. $(B \Rightarrow A) \supset (B \Rightarrow \neg C)$	4,CI,MP
6. $\neg(B \Rightarrow \neg C) \supset \neg(B \Rightarrow A)$	5,contrapositive
7. $(\neg(A \Rightarrow \neg B) \wedge A \Rightarrow \neg C \wedge \neg(B \Rightarrow \neg C)) \supset \neg(B \Rightarrow A)$	1,4,6,Disch.

I found the next result, that $A \vee B \Rightarrow C \supset (A \Rightarrow C \vee B \Rightarrow C)$, quite difficult to prove; in fact it was at one time considered to be an axiom, rather than a derived result. Many of the results of the present section also were obtained in an effort to prove it. To prove this theorem, I first prove an intermediate result, theorem 8, and from it the somewhat cumbersome result

$$(A \vee B \Leftrightarrow A) \vee (A \vee B \Leftrightarrow B) \vee (A \vee B \Rightarrow C \equiv (A \Rightarrow C \wedge B \Rightarrow C)).$$

This last result however has the much more elegant and appealing semantic counterpart:

$$f(w, \|A \vee B\|^M) = f(w, \|A\|^M) \text{ or}$$
$$f(w, \|A \vee B\|^M) = f(w, \|B\|^M) \text{ or}$$
$$f(w, \|A \vee B\|^M) = f(w, \|A\|^M) \cup f(w, \|B\|^M).$$

Theorem 10: $(A \vee B \Rightarrow C \wedge \neg(A \Rightarrow C)) \supset A \vee B \Leftrightarrow B$

Proof:

1. $A \vee B \Rightarrow C$	Supp.
2. $A \Rightarrow A \vee B$	RI
3. $A \vee C \Leftrightarrow C$	2,1,Thm. 7
4. $\neg(A \Rightarrow C)$	Supp.
5. $A \vee B \Rightarrow \neg A$	2,1,4,Thm. 8 (Cor 2)
6. $A \vee B \Rightarrow (\neg A \supset B)$	RI
7. $A \vee B \Rightarrow (\neg A \supset B) \supset ((A \vee B \Rightarrow \neg A) \supset (A \vee B \Rightarrow B))$	CI

8. $A \vee B \Rightarrow B$ 7,6,5,PC

9. $A \vee B \Leftrightarrow B$ 8,RI,PC

10. $(A \vee B \Rightarrow C \wedge \neg(A \Rightarrow C)) \supset A \vee B \Leftrightarrow B$ 1,4,9,Disch.

Theorem 11: $(A \vee B \Leftrightarrow A) \vee (A \vee B \Leftrightarrow B) \vee (A \vee B \Rightarrow C \equiv (A \Rightarrow C \wedge B \Rightarrow C))$

Proof:

1. $(A \vee B \Rightarrow C \wedge \neg(A \Rightarrow C)) \supset A \vee B \Leftrightarrow B$ Thm. 10

2. $(A \vee B \Rightarrow C \wedge \neg(B \Rightarrow C)) \supset A \vee B \Leftrightarrow A$ Thm. 10

3. $\neg(A \vee B \Rightarrow C \supset (A \Rightarrow C \wedge B \Rightarrow C)) \supset ((A \vee B \Leftrightarrow A) \vee (A \vee B \Leftrightarrow B))$ 1,2,PC

4. $(A \vee B \Leftrightarrow A) \vee (A \vee B \Leftrightarrow B) \vee (A \vee B \Rightarrow C \equiv (A \Rightarrow C \wedge B \Rightarrow C))$ 3,CC',PC

The following corollary follows immediately from the theorem by noting that each disjunct implies it. The first two disjuncts imply it via CSO; the last by PC for one direction, and CC' for the other.

Corollary (ND): $A \vee B \Rightarrow C \supset (A \Rightarrow C \vee B \Rightarrow C)$

The next theorem and corollary are useful for dealing with transitive denials of defaults, or eliminating disjunction in the antecedent of a default.

Theorem 12: $(A \vee B \Rightarrow C \wedge \neg(A \Rightarrow \neg B)) \supset B \Rightarrow C$

Proof:

1. $\neg(A \Rightarrow B) \supset (A \Rightarrow C \supset (A \wedge \neg B) \Rightarrow C)$ CV

2. $\neg(A \Rightarrow B) \supset (\neg((A \wedge \neg B) \Rightarrow C) \supset \neg(A \Rightarrow C))$ 1,PC

3. $\neg(A \vee B \Rightarrow \neg B) \supset (\neg((A \vee B) \wedge B \Rightarrow C) \supset \neg(A \vee B \Rightarrow C))$ instance of line 2

4. $\neg(A \vee B \Rightarrow \neg B) \supset (\neg(B \Rightarrow C) \supset \neg(A \vee B \Rightarrow C))$ 3,PC

5. $(A \vee B \Rightarrow \neg B) \supset (A \Rightarrow \neg B)$ ND,PC

6. $\neg(A \Rightarrow \neg B) \supset \neg(A \vee B \Rightarrow \neg B)$ 5,contrapositive

7. $\neg(A \Rightarrow \neg B)$ Supp.

8. $\neg(A \vee B \Rightarrow \neg B)$ 7,6,MP

9. $\neg(B \Rightarrow C) \supset (A \vee B \Rightarrow C)$ 8,4,MP

10. $\neg(B \Rightarrow C)$ Supp.

11. $\neg(A \vee B \Rightarrow C)$ 10,9,MP

12. $\neg(A \Rightarrow \neg B) \wedge \neg(B \Rightarrow C) \supset \neg(A \vee B \Rightarrow C)$ 7,10,11,Disch.

13. $(A \vee B \Rightarrow C \wedge \neg(A \Rightarrow \neg B)) \supset B \Rightarrow C$ — 12,PC

Corollary: $\neg(A \Rightarrow \neg B) \wedge \neg(B \Rightarrow C) \supset \neg(A \vee B \Rightarrow C)$

Theorem 13: $(\neg(A \Rightarrow \neg B) \wedge \neg(B \Rightarrow \neg C) \wedge A \Rightarrow \neg C) \supset C \Rightarrow \neg A$

Proof:

1. $\neg(A \Rightarrow \neg B) \wedge \neg(B \Rightarrow \neg C)$ — Supp.
2. $(\neg(A \Rightarrow \neg B) \wedge \neg(B \Rightarrow \neg C)) \supset \neg(A \vee B \Rightarrow \neg C)$ — Thm 12
3. $\neg(A \vee B \Rightarrow \neg C)$ — 2,1,MP
4. $A \Rightarrow \neg C$ — Supp.
5. $A \Rightarrow \neg C \supset (\neg(A \vee B \Rightarrow \neg C) \supset C \Rightarrow \neg A)$ — CP
6. $C \Rightarrow \neg A$ — 5,4,3,MP
7. $\neg(A \Rightarrow \neg B) \wedge \neg(B \Rightarrow \neg C) \wedge A \Rightarrow \neg C) \supset C \Rightarrow \neg A$ — 1,4,6,Disch

Theorem 14: $(A \Rightarrow B \wedge B \Rightarrow C \wedge \neg(A \Rightarrow C)) \supset C \Rightarrow \neg A$

Proof:

1. $A \Rightarrow B \wedge B \Rightarrow C$ — Supp.
2. $(A \vee C) \Leftrightarrow C$ — 1,Thm 7
3. $(A \vee C) \Leftrightarrow C \supset (\neg(A \Rightarrow C) \supset C \Rightarrow \neg A)$ — Thm 5
4. $\neg(A \Rightarrow C)$ — Supp.
5. $C \Rightarrow \neg A$ — 2,4,3,MP

Theorem 15: $A \Rightarrow C \supset (A \wedge B \Rightarrow C \vee A \wedge \neg B \Rightarrow C)$

Proof:

1. $\neg(A \Rightarrow B) \supset (A \Rightarrow C \supset A \wedge \neg B \Rightarrow C)$ — CV
2. $A \Rightarrow B \supset \neg(A \Rightarrow \neg B)$ — CEM (kinda)
3. $\neg(A \Rightarrow \neg B) \supset (A \Rightarrow C \supset A \wedge B \Rightarrow C)$ — CV
4. $A \Rightarrow B \supset (A \Rightarrow C \supset A \wedge B \Rightarrow C)$ — 2,3,PC
5. $A \Rightarrow B \vee \neg(A \Rightarrow B)$ — PC
6. $A \Rightarrow C \supset (A \wedge B \Rightarrow C \vee A \wedge \neg B \Rightarrow C)$ — 5,1,4,PC

The remaining results are fairly straightforward, and not crucial for later results. As a result I simply indicate how they may be proved.

Theorem 16: $((A \wedge B) \Rightarrow C \wedge (A \wedge \neg B) \Rightarrow C) \supset A \Rightarrow C$

Proof: Easy consequence of CC'.

Theorem 17: $(A \Rightarrow B \wedge B \Rightarrow C) \supset A \vee B \Rightarrow C$

Proof: Obtained from RT via PC, substituting $A \vee B$ for A in the original statement of the axiom.

Theorem 18: $A \vee B \Rightarrow C \supset (A \Rightarrow C \vee C \Rightarrow \neg A)$

Proof: Since $A \Rightarrow A \vee B$, the result follows from theorem 12. That is, $A \Rightarrow A \vee B$, $A \vee B \Rightarrow C$, $\neg(A \Rightarrow C)$ yields $C \Rightarrow \neg A$.

Theorem 19: $A \vee B \Rightarrow C \supset A \vee C \Leftrightarrow C$

Proof: Immediate from $A \Rightarrow A \vee B$ and Theorem 7.

Theorem 20: $A \vee B \Rightarrow \neg B \supset A \Rightarrow \neg B$

Proof: An instance of CV is $A \vee B \Rightarrow \neg B \supset (\neg(A \vee B \Rightarrow \neg A) \supset A \Rightarrow \neg B)$, which is $A \vee B \Rightarrow \neg B \supset (A \Rightarrow \neg B \vee (A \vee B) \Rightarrow \neg A)$. Since not $(A \vee B \Rightarrow \neg(A \wedge B))$ (assuming that $A \vee B$ is possible), the antecedent is inconsistent with the second disjunct; hence the result.

Theorem 21: $A \vee B \Rightarrow \neg B \supset A \vee B \Rightarrow A$

Proof: Follows from CI, using $A \vee B \Rightarrow (\neg B \supset A)$.

3. The First-Order Logics N and N+

The first-order language L for the variable conditional has the following primitive symbols: a denumerably infinite set of *individual variables* $x, y, z, \cdots$; a denumerably infinite set of *individual constants* $a, b, c, \cdots$; for each $n \geq 1$ a denumerably infinite set of n-place *predicate symbols*, $P, Q, R, \cdots$; together with commas, parenthesis, and the symbols $\neg, \supset, \Rightarrow$, and for. The set of individual variables and constants will make up the set of *terms*. Where no confusion arises, lower-case words may be used to stand for constants and capitalised words may be used to stand for predicate symbols. The symbols $\alpha, \beta, \gamma, \cdots$ will stand for arbitrary well-formed formulae of L.

Definition: The *sentences* or *well-formed formulae (wffs)* of L comprise the least set such that:

(i) If P is a n-place predicate symbol and $t_1, \ldots, t_n$ are terms, then $P(t_1, \ldots, t_n)$ is a wff.

(ii) If α, β are wffs then $\neg\alpha$, $(\alpha \supset \beta)$, $(\alpha \Rightarrow \beta)$, are wffs.

(iii) If α is a wff and x is an individual variable, then $(\forall x)\alpha$ is a wff.

As usual, conjunction ($\wedge$), disjunction ($\vee$), biconditionality ($\equiv$), and the existential quantifier (E) are introduced by definition. Parenthesis may be omitted where no confusion arises. Also the standard notions of scope and of free and bound variables are assumed. I will use the notation of [Mendelson 64], of using $\alpha(x)$ to indicate that α may have x as a free variable; $\alpha(t)$ then is the result of uniformly substituting term t for all free occurrences of x.

3.1. A Formal Semantics and Proof Theory

The results of this section extend the previous results of the last section to a first-order system. It proves to be the case that the extension of $NP+$ to $N+$ is exactly the same as that of NP to N. Consequently for this section I will simply deal with the weaker system N.

Sentences of L are interpreted in terms of an *N-model* $M = \langle W, f, D, V \rangle$ where W and f are as in NP, D is a domain of individuals, and V is a function on terms and predicate symbols so that

1. for term t, $V(t) \in D$,
2. for any n-place predicate symbol P, $V(P)$ is a set of $(n+1)$-tuples $\langle t_1, \ldots, t_n, w \rangle$ where each $t_i \in D$ and $w \in W$.

Given a model $M = \langle W, f, D, V \rangle$, truth at a world w is given by:

Definition:

(i) For n-place predicate symbol P, terms $t_1, \ldots, t_n$, and $w \in W$, $\vdash_w^M P(t_1, \ldots, t_n)$ iff $\langle V(t_1), \ldots, V(t_n), w \rangle \in V(P)$.

(ii) $\vdash_w^M (\forall x)\alpha$ iff for every V' which is the same as V except possibly $V(x) \neq V'(x)$, and where $M' = \langle W, f, D, V' \rangle$, $\vdash_w^{M'} \alpha$.

(iii) $\models^{M}_{w} \neg\alpha$, $\models^{M}_{w} \alpha\supset\beta$, and $\models^{M}_{w} \alpha\Rightarrow\beta$ are defined as in the propositional case.

Definitions of satisfiability and validity are analogous to those in the propositional case.

It is worth noting that the assignment of terms to individuals is fixed across possible worlds, and so we don't allow individuals to "appear" or "disappear" as alternative states of affairs are considered. Such an assumption appears reasonable given our original motivation; moreover it considerably simplifys the technical details following.

The first-order logic N is the smallest logic closed under the following axiom schemata and rules of inference:

NP: If $\vdash_{NP} \alpha$ and α' results from α by uniformly replacing each propositional variable of α by some wff of L, then α' is an axiom of N.

for-1: $(\forall x)\alpha(x)\supset\alpha(t)$ if $\alpha(x)$ is a wff of L and t is a term of L free for x in $\alpha(x)$.

for-2: $(\forall x)(\alpha\supset\beta)\supset(\alpha\supset(\forall x)\beta)$ if α is a wff of L containing no free occurrences of x.

for-3: $(\forall x)(\alpha\Rightarrow\beta)\supset(\alpha\Rightarrow(\forall x)\beta)$ if α is a wff of L containing no free occurrences of x.

The rules of inference are MP and RCM from before, together with universal generalisation:

UG: From α infer $(\forall x)\alpha$.

Definitions of theoremhood, derivability, consistency, etc., are analogous to those of NP.

For-3 has an obvious syntactic similarity to for-2. Essentially it states that for some α (with proviso), if for every individual x, $\alpha\Rightarrow\beta(x)$, then $\alpha\Rightarrow(\forall x)\beta$. This in turn effectively restricts the set of individuals in a world accessible (i.e. in $f(w, \|\alpha\|^{M})$) from some given world to be non-increasing. It is not surprising then that the Barcan formula is a theorem of N. That is, we have:

Theorem 1: $(\forall x)(\neg\alpha\Rightarrow\alpha)\supset(\neg(\forall x)\alpha\Rightarrow(\forall x)\alpha)$.

Proof:

1. $(\forall x)(\neg\alpha\Rightarrow\alpha)$	Supposition
2. $(\neg\alpha(x_i)\Rightarrow\alpha(x_i))$	1,for-1,MP
3. $(\neg\alpha(x_i)\Rightarrow\alpha(x_i))\supset(\neg(\forall x)\alpha\Rightarrow\alpha(x_i))$	Thm. 2iv, Sec. 2.4
4. $\neg(\forall x)\alpha\Rightarrow\alpha(x_i)$	2,3,MP
5. $(\forall y)(\neg(\forall x)\alpha\Rightarrow\alpha(y))$	UG
6. $\neg(\forall x)\alpha\Rightarrow(\forall y)\alpha$	5,for-3,PC

From our definition of $\Box$, this is equivalent to $(\forall x)\Box\alpha\supset\Box(\forall x)\alpha$. The converse of the Barcan formula, which essentially restricts sets of individuals in accessible worlds to be non-decreasing, is easily shown to also be a theorem of N.

3.2. Soundness and Completeness

The soundness of N is shown by the standard technique of proving the validity of the axiom schemata, and demonstrating that the rules of inference preserve validity. In brief we have:

Theorem 2: If $\vdash \alpha$ then $\models \alpha$.

Proof: Schemata NP is proven by a fairly straightforward inductive argument. Schemata for-1 and for-2 are shown to be valid by adaptations of the proofs for non-modal FOL (see, for example, [Hughes and Cresswell 68]). MP and UG are easily shown to preserve validity, and RCM follows by a proof analogous to the propositional case.

For for-3, the proof is as follows:

$\models^{M}_{w} (\forall x)(\alpha \Rightarrow \beta)$ iff for every V' like V, except possibly $V(x) \neq V'(x)$, and where $M' = \langle W, f, D, V' \rangle$, $\models^{M'}_{w} \alpha \Rightarrow \beta$.

By definition $\models^{M'}_{w} \alpha \Rightarrow \beta$ iff for $w_1 \in f(w, \|\alpha\|^{M'})$, $w_1 \in \|\beta\|^{M'}$.

However, since x is not free in α, α must have the same value at w_1 in M' as it does in M.

This means that if $\models^{M}_{w} (\forall x)(\alpha \Rightarrow \beta)$ then if $w_1 \in f(w, \|\alpha\|^{M})$ then for every V' the same as V, except possibly $V(x) \neq V'(x)$, $w_1 \in \|\beta\|^{M'}$,

or by the definition of V for for, $w_1 \in \|(\forall x)\beta\|^{M}$.

So $\models^{M}_{w} \alpha \Rightarrow (\forall x)\beta$.

This completes the proof for for-3 and hence, by the definition of N, every theorem is valid.

The completeness proof of N is an adaptation of the method of canonical models in first-order modal logics [Hughes and Cresswell 84], but modified to accommodate the variable conditional operator.

Definition: A set Γ of wffs of N has the *for-property* iff, for every wff $\alpha(x)$ and every variable x, if $\Gamma \vdash \alpha(y)$ for every variable y then $\Gamma \vdash (\forall x)\alpha$.

The following two lemmas appear in [Hughes and Cresswell 84] as theorems 9.4 and 9.5; the proofs for N are simple modifications of the proofs given therein.

Lemma 1: If Λ is a set of wffs of N that has the for-property and α is a wff of N such that $\Lambda \cup \{\alpha\}$ is consistent, then there is a maximal consistent set of wffs that includes $\Lambda \cup \{\alpha\}$ and has the for-property.

Lemma 2: If α is a consistent wff of N, then there is some maximal consistent set with the for-property that contains α.

We also have:

Lemma 3: Let Γ be a maximal consistent set of wffs with the for-property. Then for any α, where α contains no free occurrences of any variable x occurring in γ, $\Lambda = \{\gamma \mid \alpha \Rightarrow \gamma \in \Gamma\}$ has the for-property.

Proof: Assume that $\Lambda \vdash \beta(y)$ for every variable y.

Then, for every y there are $\beta_1, \ldots, \beta_n \in \Lambda$ such that $\vdash (\beta_1 \wedge \cdots \wedge \beta_n) \supset \beta(y)$.

So $\vdash (\alpha \Rightarrow \beta_1 \wedge \cdots \wedge \alpha \Rightarrow \beta_n) \supset \alpha \Rightarrow \beta(y)$ by RCK.

But $\alpha\Rightarrow\beta_1, \ldots, \alpha\Rightarrow\beta_n \in \Gamma$, since each $\beta_i \in \Lambda$.

So, $\alpha\Rightarrow\beta(y) \in \Gamma$, or $\Gamma \vdash \alpha\Rightarrow\beta(y)$ by Lindenbaum's lemma.

Since this holds for every y, and Γ has the for-property, $\Gamma \vdash (\forall x)(\alpha\Rightarrow\beta)$ for x not free in α.

By for-3, $\Gamma \vdash \alpha\Rightarrow(\forall x)\beta$.

So $(\forall x)\beta \in \Lambda$ and so $\Lambda \vdash (\forall x)\beta$, and so Λ has the for-property.

The definition of the canonical model for N follows that of NP.

Definition: The *canonical model* $<W, f, D, V>$ for N is defined as follows:

(i) $W = \{w \mid w$ is a maximal N-consistent set of wffs with the for-property$\}$.

(ii) $f(w, |\alpha|_N) = \{w_1 \in W \mid \{\beta \mid \alpha\Rightarrow\beta \in w\} \subseteq w_1\}$.

(iii) D is the set of individual variables.

(iv) For a variable x, $V(x) = x$; for a constant a, $V(a) = x$ for some variable x. For n-place predicate symbol P, terms $t_1, \ldots, t_n$, and $w \in W$, $<t_1, \ldots, t_n, w> \in V(P)$ iff $P(t_1, \ldots, t_n) \in w$.

Theorem 3: The canonical model for N is a N-model.

Proof: Analogous to that for NP-models.

Theorem 4: Let $M = <W, f, D, V>$ be a canonical model for N. Then for every sentence α and world w in M, $\vdash_w^M \alpha$ iff $\alpha \in w$.

Proof: Proof is by induction on the complexity of a formula.

As the base case, if α is a predicate symbol applied to terms, say $P(t_1, \ldots, t_n)$, then for $w \in W$, $V(P(t_1, \ldots, t_n), w) = 1$ iff $<V(t_1), \ldots, V(t_n), w> \in V(P)$

iff $<t_1, \ldots, t_n, w> \in V(P)$

iff $P(t_1, \ldots, t_n) \in w$.

The proof for $\neg$ and $\supset$ is as for NP.

For $\Rightarrow$, the *if* part is as for NP.

The *only-if* part is similar to that of NP:

Assume that $\alpha\Rightarrow\beta \notin w$.

Then $\neg(\alpha\Rightarrow\beta) \in w$ since w is maximal consistent.

So $\{\gamma \mid \alpha\Rightarrow\gamma \in w\} \cup \{\neg\beta\}$ is consistent by lemma 1, section 2.3.

So $\{\gamma \mid \alpha\Rightarrow\gamma \in w\}$ where α has no free occurrences of x in γ has the for-property by lemma 3.

By lemma 1, $\{\gamma \mid \alpha\Rightarrow\gamma \in w\} \cup \{\neg\beta\}$ has a maximal extension w_1 that has the for-property.

Since $\{\gamma \mid \alpha\Rightarrow\gamma \in w\} \subseteq w_1$, $w_1 \in f(w, |\alpha|_N)$.

Since $\neg\beta \in w_1$, $\beta \notin w_1$, and $w_1 \notin |\beta|_N$,

and so not $\vdash_w^M \alpha\Rightarrow\beta$, by the definition of $\vdash_w^M$.

For for, consider first where $(\forall x)\alpha \in w$.

By for-1, $\alpha(y) \in w$ for every variable y, and $\vdash_w^M \alpha(y)$ by the induction hypothesis.

For V' exactly like V, except that $V'(x)=V(y)$, $\vdash_w^{M'} \alpha(x)$.

So, $\vdash_w^{M'} \alpha(x)$ for all such V', and so, by the definition of V,

$\vdash_w^{M} (\forall x)\alpha$.

Conversely, assume that $(\forall x)\alpha \notin w$.

So there must be some y so that $\alpha(y) \notin w$, since w has the for-property.

So not $\vdash_w^{M} \alpha(y)$.

If V' in M' is exactly like V, except that $V'(x)=V(y)$, then not $\vdash_w^{M'} \alpha$,

and so not $\vdash_w^{M} (\forall x)\alpha$.

This completes the case for for, and so, by induction, the theorem.

Corollary: If $\vdash \alpha$ then $\vdash\!\!\!- \alpha$.

Proof: From lemma 2, every consistent wff is in some maximal consistent set with the for-property. So, by the above theorem, every consistent wff is true in some world in the canonical model. Similarly, every non-theorem is false in some world in the canonical model. Since the canonical model is an N-model, we obtain the desired result.

4. Default Reasoning

The approaches developed here for default reasoning use the logics $NP+$ and $N+$ (depending on whether one is concerned with propositional or first-order default reasoning) as the underlying formal systems. The formulation for the first approach then is slightly modified, to maintain the equivalence of the two approaches. The use of the stronger systems ($NP+$ and $N+$) also allows for simpler and more elegant equivalence proofs.

4.1. Initial Considerations

A *default theory* T is an ordered pair $<D, C>$ where D is a set of wffs of N and C is a non-empty consistent set of wffs of FOL. D is intended to represent necessary or conditional sentences constraining how the world *must* be or *could* be, while C is a set of contingent sentences constraining how the world being modelled *is*. Thus in D we would include statements such as "all ravens must be birds" and "all ravens are normally black". D may also include denials of conditional sentences, for example, "it is not the case that Quakers are normally vegetarians". Included in C would be statements such as "Opus is a raven" and "everyone taking CMPT882 this semester is under 2 metres tall". The goal is to define a "default" provability operator which, following [McDermott and Doyle 80], I will write as $T \vdash \rho$ to indicate that ρ follows by default from T.

The first part of this enterprise is surprisingly easy. Consider for example where all that is known is $Bird(opus) \Rightarrow Fly(opus)$ and $Bird(opus)$. As argued, we should not be able to conclude from this that $Fly(opus)$, simply because, while the truth of $Bird(opus)$ relies on *this* state of affairs, the truth of $Bird(opus) \Rightarrow Fly(opus)$ relies on *other*, less exceptional states of affairs, and there is no necessary connection between this state of affairs and those other states of affairs. Yet nonetheless it does seem reasonable to conclude $Fly(opus)$ "by default". The key point here is that in drawing this default conclusion, one is relying on a tacit assumption: that the world at hand is as unexceptional as possible, consistent with what is known. That is, given the above, it is entirely consistent that *opus* is a penguin, is tethered, or simply (for no known reason) does not fly. The default conclusion relies on assuming that if none of these exceptional factors are known to hold, they do not hold. This assumption can be stated as follows:

Assumption of Normality: The world being modelled is among the least exceptional worlds according to D in which the sentences of C are true.

Thus it seems that we would want to say that $T \vdash \rho$ just when, in the presence of "background information" D, ρ is true in all least exceptional worlds in which C is true[4]. Hence:

---

[4] By this I mean of course the conjunction of elements in C is true; in following formulae I will use C to stand also for the conjunction of the elements of C.

Temporary Definition: $T \vdash \rho$ iff (so it seems) $D \vdash_N C \Rightarrow \rho$.

This does in fact work in a large number of simple cases. For example if we have that

$$D = \{Raven(x) \Rightarrow Black(x), (Raven(x) \wedge Albino(x)) \Rightarrow \neg Black(x)\},^5$$

$$C = \{Raven(a)\},$$

then we can make the default conclusion that $Black(a)$. If we have that $C = \{Raven(a), Albino(a)\}$, then we can derive by default that $\neg Black(a)$; we cannot now derive $Black(a)$, because we cannot prove

$$\{Raven(x) \Rightarrow Black(x), (Raven(x) \wedge Albino(x)) \Rightarrow \neg Black(x)\} \vdash_N$$

$$(Raven(a) \wedge Albino(a)) \Rightarrow Black(a).$$

Similarly, if we have

$$D = \{Quaker(x) \Rightarrow Pacifist(x), Republican(x) \Rightarrow \neg Pacifist(x)\}$$

and learn that $Quaker(a)$, then we can conclude $Pacifist(a)$. If we learn also that $Republican(a)$, then we can conclude nothing concerning whether a, by default, satisfies $Pacifist$.

However the approach to this point also fails to work for a large class of simple cases. If we have that:

$$D = \{Raven(x) \Rightarrow Black(x)\},$$

$$C = \{Raven(a), Has\_wings(a)\}$$

then the relation

$$\{Raven(x) \Rightarrow Black(x)\} \vdash_N (Raven(a) \wedge Has\_wings(a)) \Rightarrow Black(a)$$

cannot be shown. The difficulty, from the semantic end of things, is that in all models of these sentences, in the simplest worlds where $Raven(a)$ is true, $Black(a)$ is also true. However there are models where, in the simplest worlds where $Raven(a) \wedge Has\_wings(a)$ is true, $Black(a)$ may not be true. Hence $(Raven(a) \wedge Has\_wings(a)) \Rightarrow Black(a)$ is not entailed by D.

It seems however that, based on what is known, there is no good reason for supposing that having wings has any effect on blackness. In a word, having wings seems *irrelevant* to whether a raven is black. This is the second assumption that I make in order to be able to draw default inferences in a default theory. It may be stated as:

Assumption of Relevance: Only those sentences known to bear on the truth value of a conditional relation are assumed to, in fact, have a bearing on that relation's truth value.

From these assumptions, two approaches to default reasoning are developed and, ultimately, shown to be equivalent, in that the default inferences obtained are the same in the two approaches.

---

[5] Henceforth I will, in the interests of readability, generally drop the universal quantifier. Universally quantified versions of the formulae follow trivially by universal generalisation.

4.2. The First Approach

Consider the statement $\alpha \Rightarrow \gamma$. This statement is true at a world w in model M iff $f(w, \|\alpha\|^M) \subset \|\gamma\|^M$. Intuitively, β is irrelevant to the truth of this statement if knowing β doesn't alter our judgement of the truth of the consequent of the conditional. Hence, according to our truth conditions for the conditional, β is irrelevant to $\alpha \Rightarrow \gamma$ iff $f(w, \|\alpha \wedge \beta\|^M) \subset \|\gamma\|^M$ and $f(w, \|\alpha \wedge \neg\beta\|^M) \subset \|\gamma\|^M$. So one approach is to assume, whenever possible, that a proposition β has no effect on the truth value of $\alpha \Rightarrow \gamma$. Hence, informally, we begin with a set of assertions D and extend this set by considering each conditional $\alpha \Rightarrow \gamma$ in D and each wff β of FOL, and if $\alpha \wedge \beta \Rightarrow \gamma$ is consistent, adding it to D. Thus if D is

$$\{Raven(x) \Rightarrow Black(x), (Raven(x) \wedge Albino(x)) \Rightarrow \neg Black(x)\}$$

we will add statements including

$(Raven(x) \wedge Has\_wings(x)) \Rightarrow Black(x)$ and

$(Raven(x) \wedge \neg Has\_wings(x)) \Rightarrow Black(x)$ and

$(Raven(x) \wedge Albino(x) \wedge Has\_wings(x)) \Rightarrow \neg Black(x)$.

However, this isn't quite right. If D is $\{Q(x) \Rightarrow P(x), R(x) \Rightarrow \neg P(x)\}$ we could consistently add either $(Q(x) \wedge R(x)) \Rightarrow P(x)$ or $(Q(x) \wedge R(x)) \Rightarrow \neg P(x)$ (but not both) by this recipe. The solution is to add $\alpha \wedge \beta \Rightarrow \gamma$ only if there is no other "relevant" conditional that denies γ. We begin then by identifying those conditionals that may reasonably be added to D:

Definition: $\alpha \Rightarrow \gamma$ is *supported* in Γ if there is β such that:

1. $\vdash_{FOL} \alpha \supset \beta$.
2. $\Gamma \vdash_N \beta \Rightarrow \gamma$.
3. If there is β' such that $\vdash_{FOL} \alpha \supset \beta'$ and $\Gamma \vdash_N \neg(\beta' \Rightarrow \gamma)$ then $\vdash_{FOL} \neg(\beta \supset \neg\beta')$.[6]

Thus basically $\alpha \Rightarrow \gamma$ is supported in Γ if there is a conditional $\beta \Rightarrow \gamma$ where β follows from α and there is no conditional in Γ with a stronger antecedent that denies γ. Note that it follows that if $\alpha \Rightarrow \gamma$ is supported in Γ then $\alpha \Rightarrow \gamma$ must be consistent with Γ.

For example, if $\Gamma = \{Q(x) \Rightarrow P(x), R(x) \Rightarrow \neg P(x)\}$, then $(Q(x) \wedge D(x)) \Rightarrow P(x)$ is supported in Γ since by the definition:

1. $\vdash_{FOL} (Q(x) \wedge D(x)) \supset Q(x)$ and
2. $\Gamma \vdash_N Q(x) \Rightarrow P(x)$ and
3. for no $S(x)$ where $\vdash_{FOL} Q(x) \supset S(x)$ do we have $\Gamma \vdash_N \neg(S(x) \Rightarrow P(x))$.

Alternatively, $(Q(x) \wedge D(x) \wedge R(x)) \Rightarrow P(x)$ is not supported in Γ since, while

---

[6] Note that this last condition is more general than in the approach as formulated for N in [Delgrande 88].

1. $\vdash_{FOL} (Q(x) \wedge D(x) \wedge R(x)) \supset Q(x)$ and
2. $\Gamma \vdash_N Q(x) \Rightarrow P(x)$, we also have that
3. $\vdash_{FOL} (Q(x) \wedge D(x) \wedge R(x)) \supset R(x)$ and $\Gamma \vdash_N \neg(R(x) \Rightarrow P(x))$ but not

   $\vdash_{FOL} Q(x) \supset R(x)$.

If Γ also included $(Q(x) \wedge R(x)) \Rightarrow P(x)$, then $(Q(x) \wedge D(x) \wedge R(x)) \Rightarrow P(x)$ would now be supported in Γ since, for the third condition, while $\vdash_{FOL} (Q(x) \wedge D(x) \wedge R(x)) \supset R(x)$ and $\Gamma \vdash_N \neg(R(x) \Rightarrow P(x))$ as before, we also have that $\vdash_{FOL} (Q(x) \wedge R(x)) \supset R(x)$.

As a second example, if $\Gamma = \{Q(x) \Rightarrow P(x), P(x) \Rightarrow V(x)\}$, then we will add, among other formulae, $(Q(x) \wedge P(x)) \Rightarrow V(x)$. From this we can derive, via axiom A3 of N, $Q(x) \Rightarrow V(x)$. However if Γ also contained $Q(x) \Rightarrow \neg V(x)$, then we could not add $(Q(x) \wedge P(x)) \Rightarrow V(x)$, and so could not derive $Q(x) \Rightarrow V(x)$. Hence, in this manner, plausible transitivities among defaults are included.

Using the definition of support, we can define a procedure for forming an extension to a set of defaults, where an extension is intended to account for all irrelevant properties in the defaults. If $\beta_0, \beta_1, \cdots$ is some ordering of wffs of FOL, we obtain:

Definition: A *maximal default extension* $E(D)$ of D is defined by:

1. $E_0 = D$.
2. $E_{i+1} = \delta$, where δ is defined by:

   Initially $\delta = E_i$.

   For each $D \vdash_N \alpha \Rightarrow \gamma$,

   $\delta = \delta \cup \{(\alpha \wedge \beta_i) \Rightarrow \gamma\}$ if $(\alpha \wedge \beta_i) \Rightarrow \gamma$ is supported in D;

   otherwise $\delta = \delta \cup \{(\alpha \wedge \neg\beta_i) \Rightarrow \gamma\}$.
3. $E(D) = \bigcup_{i=0}^{i=\infty} E_i$.

The procedure may be thought of as adding an inordinate number of default frame axioms to a set of defaults, in order to say that apparently irrelevant sentences are in fact irrelevant. Clearly only a single extension is produced, regardless of the ordering of wffs $\beta_0, \beta_1, \cdots$. We obtain that $D \subseteq E(D)$ and $E(E(D)) = E(D)$ for an extension. Hence, under the process of forming an extension, an extension is a *fixed point* of the set of defaults. However, if $D_1 \subseteq D_2$, it may not be the case that $E(D_1) \subseteq E(D_2)$. An example is $D_1 = \{\alpha \Rightarrow \gamma\}$ and $D_2 = D_1 \cup \{(\alpha \wedge \beta) \Rightarrow \neg\gamma\}$, wherein $E(D_1)$ contains $(\alpha \wedge \beta) \Rightarrow \gamma$ but $E(D_2)$ does not. We also obtain:

---

[6] $C \cup D$ rather than D is required if we don't have the axiom of reflexivity REF. Donald Nute and Craig Boutilier pointed this out.

Theorem 1: $E(D)$ is consistent if D is.

Proof: I will show that for E_i, if δ_j is consistent, then so is δ_{j+1}.

Assume otherwise.

So $\delta_j \cup \{\alpha \wedge \beta_i \Rightarrow \gamma\}$, and $\delta_j \cup \{\alpha \wedge \neg\beta_i \Rightarrow \gamma\}$ are both inconsistent for some β_i where

$D \vdash_N \alpha \Rightarrow \gamma$.

So for some $\{\sigma_1, \ldots, \sigma_m\} \subseteq \delta_j$, $\{\psi_1, \ldots, \psi_n\} \subseteq \delta_j$

$\{\sigma_1, \ldots, \sigma_m\} \vdash_N \neg(\alpha \wedge \beta_i \Rightarrow \gamma)$ by the compactness theorem.

So $\vdash_N (\sigma_1 \wedge \cdots \wedge \sigma_m) \supset \neg(\alpha \wedge \beta_i \Rightarrow \gamma)$,

and $\vdash_N \neg(\sigma_1 \wedge \cdots \wedge \sigma_m \wedge (\alpha \wedge \beta_i \Rightarrow \gamma))$.

Similarly, $\vdash_N \neg(\psi_1 \wedge \cdots \wedge \psi_n \wedge (\alpha \wedge \beta_i \Rightarrow \gamma))$.

So $\vdash_N \neg(\sigma_1 \wedge \cdots \wedge \sigma_m \wedge \psi_1 \wedge \cdots \wedge \psi_n \wedge (\alpha \wedge \beta_i \Rightarrow \gamma) \wedge (\alpha \wedge \neg\beta_i \Rightarrow \gamma))$

and so $\vdash_N \neg(\sigma_1 \wedge \cdots \wedge \sigma_m \wedge \psi_1 \wedge \cdots \wedge \psi_n \wedge (\alpha \Rightarrow \gamma))$ from axiom CC'.

But $\{\sigma_1, \ldots, \sigma_m, \psi_1, \ldots, \psi_n\} \subseteq \delta_j$,

and $D \subseteq \delta_j$ and $D \vdash_N \alpha \Rightarrow \gamma$,

and so δ_j is inconsistent.

By a simple inductive argument, if D is consistent, so is $E(D)$.

Corollary: If $\diamond\alpha$, then if $\alpha \Rightarrow \gamma \in E(D)$ then $\alpha \Rightarrow \neg\gamma \notin E(D)$.

Proof: If $\diamond\alpha$, then $\vdash_N (\alpha \Rightarrow \gamma) \supset \neg(\alpha \Rightarrow \neg\gamma)$.

So if $\alpha \Rightarrow \gamma \in E(D)$ then $\alpha \Rightarrow \neg\gamma \in E(D)$ would be inconsistent, contradicting the theorem above.

Theorem 2: For any $\beta \in FOL$ and $\alpha \Rightarrow \gamma \in D$ either $\alpha \wedge \beta \Rightarrow \gamma \in E(D)$ or $\alpha \wedge \neg\beta \Rightarrow \gamma \in E(D)$

Proof: Obvious from the construction of an extension.

We define default provability as we did in the last section, but now incorporating assumptions of relevance via the extension. That is:

Definition: $T \vdash \rho$ iff $E(D) \vdash_N C \Rightarrow \rho$.

Thus ρ follows by default from T if, considering all assumptions of relevance, ρ follows conditionally from the known facts C. Since $E(D)$ is consistent if D is, we obtain the easy corollary to that result:

Corollary: If $T \vdash' \rho$ then it is not the case that $T \vdash' \neg\rho$, unless T is inconsistent.

While the approach does not appear overly suited for direct implementation (at least compared to the next approach), it is worth noting that in the propositional case only a finite number of additional default statements need to be considered. Thus, since the propositional fragment of N is decidable, there is a decision procedure for default inferences by this approach.

4.3. The Second Approach:

This approach is perhaps the dual of the first. Whereas before we added assumptions to D to limit the models that we wanted to consider for a default inference, here we assume that the world at hand is among the simplest worlds, consistent with what is contingently known. So, for this approach we want to add contingent information to C. For example if we know only that

$D = \{Raven(x) \Rightarrow Black(x)\}$,

$C = \{Raven(opus), Has\_wings(opus)\}$

then if the state of affairs modelled by C were among the simplest worlds according to D then, by the definition of $\Rightarrow$, $Black(opus)$ must be true in that state of affairs. So the idea is to first make whatever conclusions we can about C under the assumption of normality. The assumption of normality is used then to add further information to C. Given such an extension (or extensions) to C we can specify that ρ follows as a default inference from T iff ρ follows in FOL from all extensions of C.

There is a minor difficulty with this approach however, arising again from the relative strength of defaults. Consider where we have:

$D = \{Raven(x) \Rightarrow Black(x), (Raven(x) \wedge Albino(x)) \Rightarrow \neg Black(x)\}$.

Thus in the least exceptional states of affairs in which there are ravens, ravens are black, and in the least exceptional states of affairs in which there are albino ravens, ravens are not black. From the semantics of $\Rightarrow$ it follows that the simplest states of affairs in which there are ravens are less exceptional than the simplest states of affairs in which there are albino ravens. This means that if we have that $C = \{Raven(opus), Albino(opus)\}$, then in extending C we should consider only the second default.

It is interesting also to see how this approach handles possible transitive relations in the defaults. Consider where we have that

$D = \{Quaker(x) \Rightarrow Pacifist(x), Pacifist(x) \Rightarrow Vegetarian(x)\}$ and $C = \{Quaker(a)\}$.

If we assume that the world at hand is among the least exceptional consistent with C, then we can conclude $Pacifist(a)$. However, given this new information, it now also becomes reasonable to conclude that $Vegetarian(a)$, barring evidence to the contrary. So effectively we need to "iterate" over default transitivities, while allowing for the fact that particular transitivities may not be warranted. Hence in the above example, if we were to add $Quaker(x) \Rightarrow \neg Vegetarian(x)$ to D, we would still want to conclude $Pacifist(a)$ but not be able to conclude $Vegetarian(a)$. This is accomplished as follows:

Definition: $\alpha \sqsupset \gamma$ is *contingently supported* in a default theory $T = \langle D, C \rangle$ if

1. $D \vdash_N \alpha \Rightarrow \gamma$.
2. $C \cup \{\alpha \sqsupset \gamma\}$ is consistent.
3. If there is α' such that $\vdash_{FOL} C \supset \alpha'$ and $D \vdash_N \neg(\alpha' \Rightarrow \gamma)$ then $\vdash_{FOL} \alpha \sqsupset \alpha'$.

This means that $\alpha \sqsupset \gamma$ is added to C, if C implies α and for any α' implied by C and which conflicts with γ, α' is implied by α. So, in a similar fashion to the first approach, only the strongest "reasonable" conditionals are contingently supported.

Note however that for condition 3, we *don't* want $\vdash_{FOL} (C \wedge \alpha) \sqsupset \alpha'$ in place of $\vdash_{FOL} C \sqsupset \alpha'$. Although the addition of α to the condition appears reasonable, and strengthens things somewhat, it has the unfortunate consequence of interfering with reasoning about contrapositives. For example, if $C = \{B\}$ and $D = \{A \Rightarrow \neg B\}$, this strengthening would block the default conclusion of $\neg A$. In addition, as will be seen, this change would also have the result that the two approaches would differ with respect to the set of default inferences obtainable.

Definition: A *maximal contingent extension* $E(C)$ of C is defined by:

1. $C_0 = C$.
2. If $\alpha \sqsupset \gamma$ is contingently supported in $\langle D, C_i \rangle$, then $C_{i+1} = C_i \cup \{\alpha \sqsupset \gamma\}$.
3. $E(C) = \bigcup_{i=0}^{i=\infty} C_i$.

Hence for $D = \{Q(x) \Rightarrow P(x), R(x) \Rightarrow \neg P(x)\}$, if C was $\{Q(x), V(x)\}$, then $Q(x) \sqsupset P(x)$ is contingently supported in $\langle D, C \rangle$ and so can be added in the formation of a maximal contingent extension of C. That is, we have that

1. $D \vdash_N Q(x) \Rightarrow P(x)$ and
2. $C \cup \{Q(x) \sqsupset P(x)\}$ is consistent and
3. there is no α' satisfying condition 3 in the definition for support.

If C was $\{Q(x), R(x)\}$ then we could no longer add $Q(x) \sqsupset P(x)$, since now condition 3 of the definition of support would be violated: we would have that $\vdash_{FOL} C \sqsupset R(x)$ and $D \vdash_N \neg(R(x) \Rightarrow P(x))$ but not $\vdash_{FOL} Q(x) \sqsupset R(x)$. However if D were augmented with $(Q(x) \wedge R(x)) \Rightarrow P(x)$, then for $C = \{Q(x), R(x)\}$ we could add $(Q(x) \wedge R(x)) \sqsupset P(x)$ to C.

If we had

$D = \{Q(x) \Rightarrow P(x), P(x) \Rightarrow V(x), R(x) \Rightarrow \neg V(x)\}$ and

$C = \{Q(x), R(x)\}$

we could add either $Q(x) \sqsupset P(x)$ or $R(x) \sqsupset \neg V(x)$ to C. In the former case we cannot (subsequently) add $P(x) \sqsupset V(x)$ since, for $P(x) \Rightarrow V(x)$ in condition 3, if α' were $R(x)$, then we would have

$\vdash_{FOL} (C \wedge (Q(x) \sqsupset P(x)) \sqsupset R(x))$ for the first part of condition 3,

and

$D \vdash_N \neg(R(x) \Rightarrow V(x))$ for the second part,

but not

$$\vdash_{FOL} P(x) \supset R(x).$$

This last example shows also that we may obtain more than one extension, and so, for this approach, we say that ρ follows by default, if ρ follows in all maximal contingent extensions.

If we use $\vdash'$ for default derivability in this approach then for default theory $T = \langle D, C \rangle$ we obtain:

Definition: $T \vdash' \rho$ iff $\cap E(C) \vdash_{FOL} \rho$.

Note that the number of extensions will typically be finite. Two extensions are distinct if and only if there are transitivities in the defaults that conflict. That is, we get more than one extension only if we have chains of defaults of the form $\alpha_1 \Rightarrow \alpha_2, \ldots, \alpha_n \Rightarrow \gamma$ along with $\beta_1 \Rightarrow \beta_2, \ldots, \beta_m \Rightarrow \neg\gamma$, where $n \geq 1$, $m \geq 2$, and where the conditionals satisfy the various conditions given in the definition of an extension. In addition, it is easily shown that in forming an extension, we need only consider those conditionals appearing explicitly in D, rather than all conditionals following from D, as is given in condition 1 for support.

Since both approaches are grounded in (arguably) sound intuitions, it would be a reasonable expectation that they give rise to the same set of default inferences. Happily, this proves to be the case and we obtain that with respect to default derivability, the two approaches are equivalent. This equivalence is shown in the next two subsections.

4.4. Approaches I and II Yield the Same Default Inferences

Notation: For maximally consistent extension $E(C)$ of C, let $\alpha_i \Rightarrow \gamma_i$ be the element of D such that $\alpha_i \supset \gamma_i$ us added to C_i to obtain C_{i+1} in the construction of $E(C)$. Thus $C_{i+1} = C_i \cup \{\alpha_i \supset \gamma_i\}$.

Theorem 3: If $\cap E(C) \vdash_N \rho$ then $E(D) \vdash_N C \Rightarrow \rho$.

Proof:

By the compactness theorem, if $E(C) \vdash_N \rho$, then for some n,

$$C \cup \{\alpha_0 \supset \gamma_0\} \cup \cdots \cup \{\alpha_n \supset \gamma_n\} \vdash_N \rho \text{ for } \alpha_i \Rightarrow \gamma_i \in D, 0 \leq i \leq n.$$

So $\vdash_N (C \wedge (\alpha_0 \supset \gamma_0) \wedge \cdots \wedge (\alpha_n \supset \gamma_n)) \supset \rho$,

and $\vdash_N (C \Rightarrow C \wedge (C \Rightarrow (\alpha_0 \supset \gamma_0)) \wedge \cdots \wedge (C \Rightarrow (\alpha_n \supset \gamma_n))) \supset C \Rightarrow \rho$ by RCK [Delgrande 87b],

or $\vdash_N ((C \Rightarrow (\alpha_0 \supset \gamma_0)) \wedge \cdots \wedge (C \Rightarrow (\alpha_n \supset \gamma_n))) \supset C \Rightarrow \rho$ by axiom ID and PC.

So $\{C \Rightarrow (\alpha_0 \supset \gamma_0), \ldots, C \Rightarrow (\alpha_n \supset \gamma_n)\} \vdash_N C \Rightarrow \rho$ by PC.

So if $(C \supset (\alpha_i \Rightarrow \gamma_i)$ for $0 \leq i \leq n$ is supported in D, then we're done.

So assume that $C \supset (\alpha_i \Rightarrow \gamma_i)$ is not supported in D.

This means that there is a β' wherein, according to the third condition of support,

1. $\vdash_{FOL} C \supset \beta'$,

2. $D \vdash_N \neg(\beta' \Rightarrow (\alpha_i \supset \gamma_i))$ and

3. not $\vdash_{FOL} Tr \supset \beta'$ where Tr is some necessarily true proposition.

From theorem 3, in section 2.5, we get that condition 2. together with the fact that $\alpha_i \Rightarrow \gamma_i$ yields $\alpha_i \Rightarrow \neg\beta'$ (i).

However, $\alpha_i \supset \gamma_i$ is conditionally supported in all extensions of C.

Hence, for every $E(C)$, we have that if $\vdash_{FOL} E(C) \supset \beta''$ where $D \vdash_N \neg(\beta'' \Rightarrow \gamma_i)$ then $\vdash_{FOL} \alpha_i \supset \beta''$ (ii).

In particular also, for the first condition we must have that $\vdash_{FOL} C \supset \beta''$.

Now since (i) and (ii) are inconsistent (assuming that α_i is not necessarily false), we must have that it is not the case that $D \vdash_N \neg(\beta' \Rightarrow \gamma_i)$ (iii).

However:

$\gamma_i \supset (\neg\alpha_i \vee \gamma_i)$ by PC.

$\beta' \Rightarrow \gamma_i \supset \beta' \Rightarrow (\neg\alpha_i \vee \gamma_i)$ from the above using RCM.

$\beta' \Rightarrow \gamma_i \supset \beta' \Rightarrow (\alpha_i \supset \gamma_i)$ from the above via PC.

$\neg(\beta' \Rightarrow (\alpha_i \supset \gamma_i)) \supset \neg(\beta' \Rightarrow \gamma_i)$ from the above via PC.

However $D \vdash_N \neg(\beta' \Rightarrow (\alpha_i \supset \gamma_i))$ and so $D \vdash_N \neg(\beta' \Rightarrow \gamma_i)$, contradicting (iii).

This completes the proof of the theorem.

The next result shows that for proofs of conditionals from $E(D)$, that negated conditionals (i.e. of the form $\neg(\alpha \Rightarrow \beta)$) are not required as premisses.

Lemma 1: If $E(D) \vdash_N C \Rightarrow \rho$ then there is a proof of $C \Rightarrow \rho$ from $E(D)$ where every premiss is of the form $\alpha \wedge \beta \Rightarrow \gamma$ where $\alpha \Rightarrow \gamma \in D$ and $\alpha \wedge \beta \Rightarrow \gamma$ is supported in D, or of the form $\alpha \supset \gamma$ for $\alpha \supset \gamma \in D$.

Proof:

A proof of $C \Rightarrow \rho$ from $E(D)$ in N consists of a series of steps, where each step is a premiss, an axiom of N, or is obtained from a previous step (or steps) by a rule of inference.

We need to show that any negated default $\neg(\alpha \Rightarrow \beta) \in D$ is expendable as a premiss in a proof.

The only place where a negated default can enter into a proof is via axioms CV and CP, since these are the only axioms (or rules) that contains a negated default.

Consider first CV: $\neg(\alpha \Rightarrow \neg\beta) \supset (\alpha \Rightarrow \gamma \supset \alpha \wedge \beta \Rightarrow \gamma)$

We have (via an inductive step) that $\alpha \Rightarrow \gamma$ is supported in D. (Note that if $D \vdash_N \alpha \Rightarrow \gamma$ then $\alpha \Rightarrow \gamma$ is supported in D.)

Assume that, $\alpha \wedge \beta \Rightarrow \gamma$ is not supported in D (otherwise we're finished).

Hence there is a β' such that $\alpha \wedge \beta \supset \beta'$, $D \vdash_N \neg(\beta' \Rightarrow \gamma)$ but it is not the case that $\neg(\alpha \Rightarrow \neg\beta')$.

So $\neg(\alpha \Rightarrow \neg\beta')$ is consistent, and since we have $\neg(\alpha \Rightarrow \neg\beta)$, we get via CV that $\alpha \wedge \beta \Rightarrow \neg\beta'$ is consistent.

However we also have that $\alpha \wedge \beta \supset \beta'$; hence $\neg(\alpha \wedge \beta \supset \neg\beta')$ contradicting the consistency of $\alpha \wedge \beta \Rightarrow \neg\beta'$.

Hence $\alpha \wedge \beta \Rightarrow \gamma$ is supported in D, and any step in a proof of $C \Rightarrow \rho$ from $E(D)$ using CV is extraneous.

Consider CP: $\alpha \Rightarrow \gamma \supset (\neg(\alpha \wedge \beta \Rightarrow \gamma) \supset \neg\gamma \Rightarrow \neg\alpha)$

Assume that $\alpha \Rightarrow \gamma$ is supported in D.

Hence $\mathrm{Tr} \Rightarrow \alpha \supset \gamma$ is supported in D where Tr is some necessarily true proposition.

Hence $\mathrm{Tr} \Rightarrow \neg\gamma \supset \neg\alpha$ is supported in D.

And so $\neg\gamma \Rightarrow \neg\gamma \supset \neg\alpha$ is supported in D.

Now we have as an instance of theorem 2.5.2 that

$$\vdash_N \neg\gamma \Rightarrow (\neg\gamma \supset \neg\alpha) \supset (\neg(\neg\gamma \Rightarrow \gamma) \supset (\neg\gamma \wedge \neg\gamma) \Rightarrow \neg\alpha).$$

If we assume that $\neg(\neg\gamma \Rightarrow \gamma)$, then we obtain $\neg\gamma \Rightarrow \neg\alpha$. (If we don't have $\neg(\neg\gamma \Rightarrow \gamma)$ then a straightforward argument shows that trivially we get $\neg\gamma \Rightarrow \neg\alpha$.)

Hence $\neg\gamma \Rightarrow \neg\alpha$ is supported in D, and CP is expendable, in the same manner that CV is.

Hence any negated default $\neg(\alpha \Rightarrow \beta) \in D$ is expendable as a premiss in a proof of $C \Rightarrow \rho$ from $E(D)$.

Lemma 2: For $E(D) \vdash_N C \Rightarrow \rho$ there is a proof of $C \Rightarrow \rho$ from $E(D)$ where every premiss $\alpha \Rightarrow \gamma$ is such that $\alpha \supset C$.

Proof:

A proof of $E(D) \vdash_N C \Rightarrow \rho$ consists of a series of steps, where each step is a premiss, an axiom of N, or is obtained from a previous step (or steps) by a rule of inference.

By examining the axioms and rule of inference RCM, the condition can be seen to hold, with the exception of CV and RC.

A straightforward inductive argument then gives the desired result, with the exception of axioms CV and RC.

However, from the previous lemma, we have that, effectively, these axioms may be replaced by the notion of support in D, hence the desired result.

Lemma 3: If $\alpha \Rightarrow \gamma \in D$, and $\alpha \wedge \beta \Rightarrow \gamma$ is supported in D and $\alpha \wedge \beta \supset C$ then $\alpha \wedge \beta \supset \gamma \in E(C)$.

Proof:

If we can show that $\alpha \supset \gamma$ is contingently supported, then clearly $\alpha \wedge \beta \supset \gamma \in E(C)$.

The first condition for contingent support is trivially satisfied. For the second, we need to show that $C \cup \{\alpha \supset \gamma\}$ is consistent. Assume otherwise:

Then $C \vdash_{FOL} \alpha \wedge \neg\gamma$, and so $C \vdash_{FOL} \neg\gamma$.

Hence $\vdash_N C \supset \neg\gamma$ and so $\vdash_N \neg(C \supset \gamma)$.

Since $\alpha \wedge \beta \Rightarrow \gamma$ is supported in D, and $\alpha \wedge \beta \supset C$ and $D \vdash_N \neg(C \Rightarrow \gamma)$, we must have that $\vdash_{FOL} \alpha \supset C$.

Hence $C \equiv \alpha$.

But this is a contradiction, since $\alpha \Rightarrow \gamma$ and $\neg(C \Rightarrow \gamma)$.

Hence $C \cup \{\alpha \supset \gamma\}$ is consistent.

For the third condition, assume that there is α' such that $\vdash_{FOL} C \supset \alpha'$ and $D \vdash_N \neg(\alpha' \Rightarrow \gamma)$.

We have that $\vdash_N \alpha \wedge \beta \supset C$, and $C \supset \alpha'$, and so $\vdash_N \alpha \wedge \beta \supset \alpha'$.

Since we also have that $D \vdash_N \neg(\alpha' \Rightarrow \gamma)$, we obtain (from the support in D of $\alpha \wedge \beta \Rightarrow \gamma$) that

$\vdash_{FOL} \alpha \supset \alpha'$.

Hence $\alpha \supset \gamma$ is contingently supported, and so $\alpha \supset \gamma \in E(C)$, and so $\alpha \wedge \beta \supset \gamma \in E(C)$.

Theorem 4: If $E(D) \vdash_N C \Rightarrow \rho$ then $\cap E(C) \vdash_{FOL} \rho$.

Proof:

If $E(D) \vdash_N C \Rightarrow \rho$ then by the compactness theorem, and the previous two lemmas, there is a finite Γ where $\Gamma \subseteq E(D)$ and where Γ contains only defaults $\alpha \wedge \beta \Rightarrow \gamma$ where $\alpha \Rightarrow \gamma \in D$ and $\alpha \wedge \beta \Rightarrow \gamma$ is supported in D, and $\alpha \wedge \beta \supset C$.

However, we know that for each $\alpha \wedge \beta \Rightarrow \gamma \in \Gamma$, $\alpha \wedge \beta \supset \gamma \in E(C)$.

Also, for any proof of $\Gamma \vdash_N C \Rightarrow \rho$, if we uniformly substitute $\supset$ for $\Rightarrow$, we obtain a proof in FOL. That is, we obtain that $\Gamma' \vdash_{FOL} C \supset \rho$ where Γ' is Γ with all occurrences of $\Rightarrow$ replaced with $\supset$. (Proof of this follows easily by induction.)

Since $E(C) \vdash_{FOL} \Gamma$ and $E(C) \vdash_{FOL} C$, we obtain

$E(C) \vdash_{FOL} \rho$ as required.

Corollary: $T \vdash \rho$ iff $T \vdash' \rho$.

5. Examples

5.1. Examples of Default Reasoning

The approach to default reasoning arguably leads to reasonable and intuitive default inferences. As a first example, assume that we have the default portion of a theory:

$$D_1 = \{Adult(x) \Rightarrow Employed(x), Univ\_st(x) \Rightarrow \neg Employed(x)\},$$

say, adults are typically employed, while university students normally are not. If we knew that someone was an adult then we could conclude by default that that individual was employed. If we knew that someone was an adult and a university student, then we could draw no conclusion. If, on the other hand, we knew that someone was an adult and was Dutch, then we would still conclude that they were employed. Of course, we also know that university students are typically adults, and so the defaults could be augmented to:

$$D_2 = D_1 \cup \{Univ\_st(x) \Rightarrow Adult(x)\}.$$

Now if we were told that someone was an adult and a university student, we *would* conclude by default that that person was not employed. The reason that we can now draw a conclusion is that in any model of D_2, in the simplest worlds in which someone is a university student, that person is not employed but is an adult. From a proof-theoretic viewpoint, the reason that we are able to draw this conclusion is that we can prove that $D_2 \vdash_N Univ\_st(x) \Rightarrow (\neg Employed(x) \wedge Adult(x))$ from which we can prove that $D_2 \vdash_N (Univ\_st(x) \wedge \neg Employed(x)) \Rightarrow Adult(x)$, from which the result follows immediately. This result is closely related to the notion of the ordering of defaults, which is examined in the next subsection.

From the logic N, we have the relation:

$$D_2 \vdash_N Adult(x) \Rightarrow \neg Univ\_st(x),$$

and so from N we can *derive* the default that, given D_2, adults are normally not university students.

The defaults in D_2 are of the form $\alpha \Rightarrow \beta$, $\beta \Rightarrow \gamma$, $\alpha \Rightarrow \neg\gamma$. In [Delgrande 87a] it was claimed, erroneously, that if α was known to be contingently true that no default conclusion concerning γ should be forthcoming. The above example hopefully illustrates that we would in fact want to draw a default conclusion, that $\neg\gamma$ is the case. The reason why this should be so is that, from the point of view of the semantics, in the simplest worlds in which α is true, $\neg\gamma$ must also be true. While it is also true that in the simplest worlds in which α is true, β must also be true, it doesn't follow that in the simplest worlds in which β is true, γ must also be true. Note however that this argument only goes through for "direct" defaults of the form $\alpha \Rightarrow \neg\gamma$. If we had instead $\alpha \Rightarrow \beta'$, $\beta' \Rightarrow \neg\gamma$, along with $\alpha \Rightarrow \beta$, $\beta \Rightarrow \gamma$, then we would draw no conclusion concerning γ.

Consider next the defaults:

$$D_3 = \{Raven(x) \Rightarrow Black(x), Raven(x) \Rightarrow Fly(x), (Raven(x) \wedge Albino(x)) \Rightarrow \neg Black(x)\}.$$

Not unexpectedly, we can conclude by default that ravens with wings are black, and that ravens that fly (or don't fly) are black. Moreover albino ravens are concluded by default to fly but to not be black. Consider further where we augment the defaults so that we have:

$$D_4 = D_3 \cup \{Bear(x) \Rightarrow Black(x), (Bear(x) \wedge Has\_Illness\_X(x)) \Rightarrow \neg Black(x)\}.$$

The default conclusions in D_3 go through as before. However, if we learn that a particular raven has illness X then we would *not* conclude by default that the raven was not black; rather we would still conclude that the individual was black. The reason for this is that, by our notion of relevance, illness X has no apparent connection with the colouring of ravens, even though it clearly does for bears.

If we want to treat albinoism as affecting all animals, then we could do so as follows, retaining our previous information concerning ravens:

$$D_5 = \{(Animal(x) \wedge Albino(x)) \Rightarrow White(x), Raven(x) \Rightarrow Black(x),$$
$$\Box(Raven(x) \supset Animal(x)), Raven(x) \Rightarrow Fly(x), \Box(White(x) \supset \neg Black(x))\}.$$[7]

From these defaults we can draw the same default conclusions concerning raven colouring that we could in D_3.

Transitive relations among the defaults appear to be handled correctly. If we have:

$$D_6 = \{Quaker(x) \Rightarrow Pacifist(x), Pacifist(x) \Rightarrow Vegetarian(x)\},$$

and we know contingently that $Quaker(a)$ and $Republican(a)$, then we could conclude by default that $Vegetarian(a)$. If we were to augment D_6 with either $Republican(x) \Rightarrow \neg Pacifist(x)$ or $Republican(x) \Rightarrow \neg Vegetarian(x)$ then in neither case could we conclude $Vegetarian(a)$, nor could we if $\neg(Republican(x) \Rightarrow Pacifist(x))$ or $\neg(Republican(x) \Rightarrow Vegitarian(x))$ were added. If, on the other hand, we have:

$$D_7 = D_6 \cup \{(Pacifist(x) \wedge Republican(x)) \Rightarrow \neg Vegetarian(x)\},$$

and we know that $Quaker(a)$, then we could conclude $Vegetarian(a)$ via the default transitivity. If we knew that $Quaker(a)$ and $Republican(a)$, then using the first default we could conclude that $Pacifist(a)$. However, since we have that pacifists are normally vegetarian, but that Republican pacifists are normally not vegetarian, we could also conclude $\neg Vegetarian(a)$.

Finally, the approach applies to "circularities" in defaults. Consider for example where we have that Quakers are normally doves; Republicans are normally hawks; both doves and hawks are normally politically inclined; but necessarily no doves are hawks:

$$D_8 = \{Quaker(x) \Rightarrow Dove(x), Republican(x) \Rightarrow Hawk(x), Dove(x) \Rightarrow Polit\_incl(x),$$

---

[7] For those concerned about such matters, the Barcan formula and its converse are theorems of the logic *N*, and so no problem arises with respect to the ordering of quantifiers and the necessity operator.

$$Hawk(x) \Rightarrow Polit\_incl(x), \Box(Dove(x) \supset \neg Hawk(x))\}.$$

Such a set of defaults falls outside the scope of the approaches of, for example, [Touretzky 84] and [Horty et al 87], since there is a "circularity" between *Dove* and *Hawk*. In the graphical representation of these approaches there would be a cycle of links between *Dove* and *Hawk* and consequently the statements of D could not be represented as a directed acyclic graph, as these approaches require.

In the approach at hand though, if we know that $Quaker(a)$ then we can conclude that normally $Dove(a)$ (and hence $\neg Hawk(a)$) and subsequently $Polit\_incl(x)$. If we knew that both $Quaker(a)$ and $Republican(a)$, then we could not conclude by default either that $Dove(a)$ or $Hawk(a)$. Nonetheless we can still conclude that $Polit\_incl(a)$.

5.2. The "Ordering" of Default Statements

A major difficulty with many extant systems of default reasoning is that a default theory may give rise to more than one *extension*, or distinct set of beliefs that may consistently be held. To address this problem, a number of schemes have been proposed for use in specialised circumstances – notably for inheritance of properties – for ordering defaults so that only a single extension is produced. Such schemes, which may be intuitively appealing and which may yield a single, apparently reasonable extension, are nonetheless lacking in any firm formal foundation. Since the approach at hand produces only a single extension for a default theory, it is an interesting question to ask, first, how "orderings" of defaults are handled in the approach and, second, how the semantics of the approach at hand explicates other default-ordering schemes.

Consider the first example of the previous section:

$$D = \{Adult(x) \Rightarrow Employed(x), Univ\_st(x) \Rightarrow \neg Employed(x), Univ\_st(x) \Rightarrow Adult(x)\}.$$

If we knew that *sue* was both an *Adult* and *Univ_st*, then we could conclude by default that *sue* was not employed. So the second default is "applied" over the first. If the default theory lacked the third default statement, then we could not draw this conclusion. Why this is so can be seen from an examination of the semantic theory. For a statement $\alpha \Rightarrow \beta$, we have that $f(w, \|\alpha\|^M) \subseteq \|\beta\|^M$. What this means is that the least exceptional worlds in which α is true (in M) also have β true. Hence (and this is the key) the least exceptional worlds in which β is true are no more exceptional than the least exceptional worlds in which α is true. We can define a partial order then on the sets of simplest worlds in which propositions are true as follows:

Definition: If $\models^M_{\sim} \alpha \Rightarrow \beta$ then $f(w, \|\beta\|^M) \leq f(w, \|\alpha\|^M)$.

For the above example we obtain that

$$f(w, \|Employed(x)\|^M) \leq f(w, \|Adult(x)\|^M),$$

$$f(w, \|\neg Employed(x)\|^M) \leq f(w, \|Univ\_st(x)\|^M),$$

$$f(w, \|Adult(x)\|^M) \leq f(w, \|Univ\_st(x)\|^M).$$

Hence by transitivity

$$f(w, \|Employed(x)\|^M) \leq f(w, \|Univ\_st(x)\|^M).$$

Clearly also $f(w, \|Employed(x)\|^M)$ and $f(w, \|\neg Employed(x)\|^M)$ are disjoint and so, since $f(w, \|Univ\_st(x)\|^M) \subseteq \|Adult(x)\|^M$, $f(w, \|Adult(x)\|^M)$ is strictly less exceptional than $f(w, \|Univ\_st(x)\|^M)$. We can write this as

$$f(w, \|Adult(x)\|^M) < f(w, \|Univ\_st(x)\|^M).$$

And now the reason for concluding that *sue* is not employed is (hopefully) clear: *sue* is both an adult and a university student, and we assume that the world being modelled (call it w_R) is as unexceptional as possible consistent with what is known. Thus

$$w_R \in \|Adult(sue)\|^M \cap \|Univ\_st(sue)\|^M$$

and since $f(w_R, \|Adult(sue)\|^M) < f(w_R, \|Univ\_st(sue)\|^M)$, we obtain that

$$w_R \notin f(w_R, \|Adult(sue)\|^M).$$

But it is consistent that $w_R \in f(w_R, \|Univ\_st(sue)\|^M)$ and so, under the assumption of normality, we can conclude that in fact $w_R \in f(w_R, \|Univ\_st(sue)\|^M)$. From the second default we have that $f(w_R, \|Univ\_st(sue)\|^M) \subseteq \|\neg Employed(sue)\|^M$. Hence $w_R \in \|\neg Employed(sue)\|^M$; thus *sue* is not employed.

In a nutshell then, we can define a relation $\leq$ between sets of worlds determined by propositions and this, together with the assumption of normality, leads to an implicit partial ordering of default statements. Note that there is no question as to how defaults should be "ordered": this is dictated by the model theory (and the assumption of normality). Note also that, in contrast with other approaches, if one disagrees with such an "ordering" that there is a structure that one can examine and perhaps alter: again, the model theory.

If we equate default statements in N with default rules in Reiter's system, we can use these considerations to rationalise the ordering proposed in, for example, [Touretzky 84]. There, for default statements D_i and D_j, $D_i \ll D_j$ is defined to hold[8] if there is a default statement $\frac{\alpha : M\beta}{\beta}$ where α is the prerequisite of D_i and β is the prerequisite of D_j, or for some D_n, $D_i \ll D_n \ll D_j$. For conflicting default proofs S_k and S_l, $S_k \ll S_l$ if D_i and D_j are the maximal (wrt $\ll$) rules in the proofs of S_k and S_l respectively and $D_i \ll D_j$. If $S_k \ll S_l$ then the proof of S_l is rejected.

[8] Touretzky uses the symbol $\leq$.

Consider this in terms of the logic N: If $S_k \ll S_l$ then $D_i \ll D_j$ for the maximum respective rules. Hence there is a rule

$$\frac{Prereq(D_i) : M\ Prereq(D_j)}{Prereq(D_j)},$$

or correspondingly in N,

$$Prereq(D_i) \Rightarrow Prereq(D_j),$$

or

$$f(w, \|Prereq(D_j)\|^N) < f(w, \|Prereq(D_i)\|^N).$$

(We obtain < rather than ≤ since S_k and S_l conflict.) By the argument at the start of this section, under the assumption of normality the real world may be one of those in $f(w, \|Prereq(D_i)\|^N)$ but cannot be one of those in $f(w, \|Prereq(D_j)\|^N)$ (although it is in $\|Prereq(D_j)\|^N$). Hence the proof with D_i should be adopted; hence S_k and not S_l is "selected".

The scheme proposed by [Horty et al 87] may be similarly rationalised by appeal to the model theory of N. The present approach also extends these approaches in that it can deal with "circularities" in the defaults (as illustrated in the last example of the previous subsection) and can deal with denials of defaults.

5.3. The Modified Yale Shooting Problem

While the approach at hand appears to handle conflicting defaults correctly, [Hanks and McDermott 86] shows that there are situations wherein there are no (apparently) conflicting defaults, yet multiple extensions nevertheless arise. Their paper provides a simple example in commonsense reasoning and shows that the major approaches to default reasoning fail to adequately deal with it. In their example a gun is loaded; some time passes; and a person is shot. Assuming that the person is initially alive, a reasonable default conclusion is that the person is no longer alive after being shot. Yet the systems of Reiter, McDermott and Doyle, and McCarthy are all shown to yield two extensions – one in which the person is not alive, and one in which he is.[9] For example, the straightforward encoding of this example in the situation calculus so that circumscription may be applied is as follows:

$$T(Alive, s_0),$$

$$\forall s\ T(Loaded, Result(Load, s)),$$

---

[9] ··· or she is. At any rate, the obvious representation of this example in these formalisms is shown to yield two possible outcomes. This is not to say though that with sufficient ingenuity the example could not be otherwise represented so that multiple extensions do not occur. Yet the authors argue persuasively that this is not the case and the problem of multiple extensions is intrinsic to these formalisms.

$\forall s\ T(Loaded, s) \supset (Ab(Alive, Shoot, s) \wedge T(Dead, Result(Shoot, s)))$,

$\forall f, e, s\ (T(f, s) \wedge \neg Ab(f, e, s)) \supset T(f, Result(e, s))$.

$Ab(f, e, s)$ is intended to capture the notion that fact f is abnormal with respect to event e in state s (hence being *Alive* is abnormal with respect to *Shoot*). $T(f, s)$ states that fact f is true in state s. $Result(e, s)$ is a function whose value is the state resulting from event e occurring in state s. As mentioned, circumscribing Ab (to minimise abnormal happenings) yields two conflicting extensions – one where the person is alive and one where he is dead.

The obvious representation of this example in N is:

$T(Alive, s_0)$,

$\forall s\ T(Loaded, Result(Load, s))$,

$\forall s\ T(Loaded, s) \Rightarrow T(Dead, Result(Shoot, s))$,

$\forall f, e, s\ T(f, s) \Rightarrow T(f, Result(e, s))$.

Instead of an abnormality predicate we use the conditional operator. The third statement may be read as "if the gun is loaded in s then in the normal course of events the person will be dead in the state resulting from a *Shoot* event in s". The fourth statement says that things normally remain the same over an event.

Given this formulation, it is quite *consistent* that the person may be alive or dead in the final state – as indeed it should be. However the approach to default reasoning based on N yields *no* conclusion as to the person's final state of health. If $s_2 = Result(Wait, Result(Load, s_0))$, we can derive $T(Loaded, s_2)$ and $T(Alive, s_2)$ as expected. Thus after a *Load* and a *Wait*, the gun is loaded and the person alive. However from the third statement we obtain

$T(Loaded, s_2) \Rightarrow T(Dead, Result(Shoot, s_2))$,

and from the fourth statement we obtain

$T(Alive, s_2) \Rightarrow T(Alive, Result(Shoot, s_2))$.

If we assume that *Dead* persons are not *Alive*, then the above two statements conflict and so the problem in this formulation *does* result from conflicting defaults. Hence, quite reasonably, no conclusion is forthcoming.

However in this case we do want to draw a conclusion, namely that the person is *Dead*. The problem arises from the fourth statement of the formulation, where as an instance we obtain:

$\forall e, s\ T(Alive, s) \Rightarrow T(Alive, Result(e, s))$.

That is, people tend to remain alive over all events. This though is quite simply false: people tend to not remain alive if, for instance, they are shot. Hence the last statement could be replaced most perspicuously by:[10]

[10] Assuming that equality is introduced into the logic.

$\forall f, e, s\ (f = Alive) \vee (T(f, s) \Rightarrow T(f, Result(e, s)))$,

$\forall e, s\ (e = Shoot) \vee (T(Alive, s) \Rightarrow T(Alive, Result(e, s)))$.

In this case we obtain the desired default inference $T(Dead, Result(Shoot, s_2))$.

An alternative, if we have conflicting defaults, is to simply state what happens in the case of a conflict. Thus if we add

$\forall s\ ((T(Alive, s) \wedge T(Loaded, s)) \Rightarrow T(Dead, Result(Shoot, s)))$

to the original formulation we again obtain the inference $T(Dead, Result(Shoot, s_2))$. This alternative amounts to saying that things tend to remain the same (sentence 4 in the original formulation) unless explicitly overridden. Finally, note that the example of conflicting defaults concerning Suc's employment status, provides yet another way of resolving conflicting defaults.

6. Conclusion

See the previous papers for a reasonable set of conclusions regarding claims concerning the use, efficacy, and general all-round utility of the logics and approaches to default reasoning. As mentioned in the introduction, this report is intended to cover all aspects of the logics, with respect to the formal systems themselves and the approaches to default reasoning; if there is anything omitted, unclear or just plain wrong, please let me know.

Extant work concerning complexity issues, and yet another (but nice!) approach to default reasoning, are not included here, but rather are the subject of a planned-for second technical report containing yet more interesting and somewhat less-interesting proofs on these various subjects. This second report will also (presumably) contain proofs extending the results contained herein to the full set of normal conditional logics.

Acknowledgements

I would like to thank Sharon Hamilton, Howard Hamilton, Robert Hadley, Craig Boutillier, and Ken Jackson for their comments on various aspects of this work. This research was supported by the Natural Science and Engineering Research Council of Canada grant A0884.

References

[1] B.F. Chellas, "Basic Conditional Logic", *Journal of Philosophical Logic 4*, 1975, pp 133-153.

[2] J.P. Delgrande, "A Propositional Logic for Natural Kinds", *AI-86 Canadian Society for Computational Studies of Artificial Intelligence Conference*, May 1986a.

[3] J.P. Delgrande, "A First-Order Logic for Prototypical Properties: Extended Report", Technical Report 86-15, Laboratory for Computer and Communications Research, School of Computing Science, Simon Fraser University, 1986b.

[4] J.P. Delgrande, "An Approach to Default Reasoning Based on a First-Order Conditional Logic", *Proc. AAAI-87*, Seattle, 1987a, pp. 340-345.

[5] J.P. Delgrande, "A First-Order Logic for Prototypical Properties", *Artificial Intelligence*, 33, 1, 1987b, pp. 105-130.

[6] J.P. Delgrande, "An Approach to Default Reasoning Based on a First-Order Conditional Logic – Extended and Revised Report", LCCR TR 87-9, School of Computing Science, Simon Fraser University, 1987c.

[7] J.P. Delgrande, "An Approach to Default Reasoning Based on a First-Order Conditional Logic: Revised Report", *Artificial Intelligence Journal, 36, 1*, August 1988, pp. 63-90.

[8] S. Hanks and D. McDermott, "Default Reasoning, Nonmonotonic Logics, and the Frame Problem", *Proc. AAAI-86*, Philadelphia, 1986, pp 328-333.

[9] J.F. Horty, R.H. Thomason, and D.S. Touretzky, "A Skeptical Theory of Inheritance in Nonmonotonic Semantic Networks", *Proc. AAAI-87*, Seattle, 1987, pp 358-363.

[10] G.E. Hughes and M.J. Cresswell, *An Introduction to Modal Logic*, Methuen and Co. Ltd., 1968.

[11] G.E. Hughes and M.J. Cresswell, *A Companion to Modal Logic*, Methuen and Co. Ltd., 1984.

[12] D. Lewis, *Counterfactuals*, Harvard University Press, 1973.

[13] D. McDermott and J. Doyle, "Non-Monotonic Logic I", *Artificial Intelligence 13*, 1980, pp 41-72.

[14] E. Mendelson, *Introduction to Mathematical Logic*, D. Van Nostrand Co., 1964.

[15] D. Nute, *Topics in Conditional Logic*, Philosophical Studies Series in Philosophy, Volume 20, D. Reidel Pub. Co., 1980.

[16] R.F. Stalnaker, "A Theory of Conditionals", in *Studies in Logical Theory*, N. Rescher (ed.), Basil Blackwell, Oxford, 1968, pp 98-112.

[17] D.S. Touretzky, "Implicit Ordering of Defaults in Inheritance Systems", *Proc. AAAI-84*, Austin, 1984, pp 322-325.

Centre for Systems Science

• Simon Fraser University established the **Centre for Systems Science** as a provincial centre for advanced study. The objective of CSS is twofold: coordinate and combine advanced systems research at SFU, and encourage application of research results within industry in order to further the technological development of British Columbia.

•CSS has assembled three research groups to streamline its efforts. The orientation of these groups is as follows: intelligent systems, microelectronics, and computer and communications systems. Because of shared research interests, this latter group merges with LCCR (**Laboratory for Computer and Communications Research**) for the publication of technical reports.

• CSS's three groups involve over 80 skilled faculty researchers principally from SFU's Schools of Engineering Science and Computing Science, but also draw substantially on Mathematics, Physics, Geography, Philosophy, Linguistics, Psychology, Communication, and Kinesiology.